

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW nucleic - nucleic search, using sw model

Run on: June 1, 2003, 23:40:52 ; Search time 292 Seconds  
(without alignments)  
7288.145 Million ccll updates/sec

Title: US-09-978-274A-1  
Perfect score: 945  
Sequence: 1 atgaaggtatctgttagt.....atctattgaaggattctaa 945

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Capext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_101002.\*  
1: /SID52/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SID52/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SID52/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SID52/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SID52/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SID52/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SID52/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SID52/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SID52/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SID52/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SID52/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SID52/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SID52/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SID52/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SID52/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SID52/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SID52/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SID52/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SID52/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SID52/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SID52/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SID52/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	632.4	66.9	1379	21	AAZ59220 pokeweed antiviral
2	632.4	66.9	1379	21	AAZ45197 Wild-type pokeweed
3	630.8	66.8	1195	15	AAQ56672 Sequence of Phytol
4	630.8	66.8	1379	22	AAC87929 P. americana pokew
5	629.2	66.6	1195	16	AAQ81457 Phytolacca antivir
6	629.2	66.6	1379	21	AAZ59221 Variant pokeweed a
7	624.2	66.1	2472	14	AAQ43967 Pokeweed antiviral
8	588.6	62.3	882	19	AAAT99556 Phytolacca insular
9	586.4	62.1	2369	19	AAAT99557 Phytolacca insular

06/02/2003

Phytolacca insular  
Antiviral protein  
DNA pJMC201 fragile  
Wild-type pokeweed  
Cloning vector pRI  
DNA of ribosome  
Saporin encoding seq  
M13mpl8-G2, a sapor  
M13mpl8-G9, a sapor  
Saporin clone M13  
Saporin clone M13  
Saporin cDNA clone  
Saporin cDNA clone  
Saporin cDNA clone  
Saporin cDNA clone  
Gelatin toxin DNA  
DNA encoding a sap  
DNA encoding chemo  
DNA encoding chemo  
DNA encoding chemo  
Saporin encoding seq  
M13mpl8-G1, a sapor  
Saporin clone M13  
Saporin cDNA clone  
VEGPI65-ATMEL-SAP  
Saporin encoding seq  
Saporin encoding seq  
M13mpl8-G4, a sapor  
M13mpl8-G2, a sapor  
Saporin clone M13  
Saporin clone M13

ALIGNMENTS

RESULT 1  
AAZ59220 standard: cDNA: 1379 BP.  
AAZ59220:

20-APR-2000 (first entry)

Pokeweed antiviral protein coding sequence spring leaf form.

Pokeweed antiviral protein spring leaf form; PAP: transgenic plant;  
Resistance; Potato virus X; potato virus Y; potato leaf roll virus;  
tuber; ss.

Phytolacca americana.

US6015940-A.

18-JAN-2000.

07-APR-1992: 92US-0865169.

07-APR-1992: 92US-0865169.

(MUNS ) MONSANTO CO.

Koniewski WK, Turner NE, Lodge JK;

WPI: 2000-126326/11.

Production of transgenic potato plants or tubers expressing pokeweed  
antiviral protein which are resistant to potato virus X or Y.



Query Match	66.9%	Score 632.4	DB 21	Length 1379
Best Local Similarity	80.0%	Pred. No. 2.1e-166		
Matches 757	Conservative 0	Mismatches 186	Indels 3	Gaps 1
QY	3	GAAGGTGATGCTTGTAGTTGTGTGTGACGTTTAATAGCGTGGCTCATTTGCTGCACCAACTTC	62	
Db	221	GAAGATGAAGTCGATCGCTTCTGTGTGACAATATCAATATGGCTCAATCTTGCACCAACTTC	280	
QY	63	AACCTGTGGCCATAAATACGATCACTCTTGTATGCTGGAAATGCCACCAATTAACAANAATATGC	122	
Db	281	AACCTGTGGCGTGAATACAAATCATCTACAATGTGGAAAGTACCAACATATAGCAAAATACGC	340	
QY	123	CACCTTTATGGAAATCTCTCGTAATCAAGGGAAGATCCAAAATAAAATGCTATGGCAT	182	
Db	341	CACCTTTCTGAATGCTCTCGTAATGAGCGAAGATCCCAAGTTTAAATGCTATGGAT	400	
QY	183	ACCAATGCTACCTGATACATAATTGACCCCTAAGTACTTATGCTTAAGCTCCAAGTGC	242	
Db	401	ACCAATGCTGCCAATACAAATACAAATCCAAATACGCTGTGGTTCAGCTCCAAGTTC	460	
QY	243	AAACCTAAAACCATTACACTAATGCTGACACGAATAACTTATACGTGATGGCTATTC	302	
Db	461	AAATAAAAACCATCACACTAATGCTGACACGAACAATTTGTATGCTGATGGTTATTC	520	
QY	303	TGATCCCTTCAATGGCAATAAGTGTGGTTACCATATATATTAATGATATTAACAAGCACCGA	362	
Db	521	TGATCCCTTGAACCAATAAATATGCTGTTACCATATCTTTAATGATATCTCAAGTACTGA	580	
QY	363	ACGCACTGATGTGGACAATACTCTTTTCCCTCAAGTCTGATGTTCTCGTGTGCAATGTCCAT	422	
Db	581	AGCCCAAGATGTAGACACTACTCTTTGCCCAATGCCNATTCGGTGTAGTAAAACAT	640	
QY	423	TAACTACAATAGCTTTATATCCGACCATTGGAAGAGAAAGACAGAAATCAACAATAATCA	482	
Db	641	AAACTTTGATAGTCGATATCCAAATTTGGAATCAAAAGCGGAGTAAAATCAAGAAGTCA	700	
QY	483	AGTCCCAATTTGGAAATTCAAATACTCAGCAGTGCATTTGGAANAATCTCTCGAGTTGATTC	542	
Db	701	GGTCCCACTGGGAATTTCAAAATCTGCACAGTATATTTGGAAAGATTTCTGAGAGTGTGTC	760	
QY	543	ATTCCCTGTAAAACATCAGCTTTTCTTCTAGCTGGTAGCCATCCAAATGGTTTTCAGAGGC	602	
Db	761	ATTCACTGAGAAACCGAAGCCGAATTTCTTATTTGGTAGCCATACAATGTATCAGAGGC	820	
QY	603	ACGCCGATTCAGTACATAGAACCAAGTCAAGACTAATTTTAATAGACCAATCTACCC	662	
Db	821	ACCAAGTTCAAGTACATAGAGAAATCAGGTGAAAATTAATTTTACAGAGCATTTCAACCC	880	
QY	663	TGATCCCAAGTAATTAATTTGGAGGAGAGTGGGCAAAATCTCTGAGGCAATTCACAA	722	
Db	881	TATCCCAAGTACTTTAATTTTGAAGAGACATGGGGTAAGNTTTCAACAGCAATTCATGA	940	
QY	723	TGCCAAGATGGGCTTTTACCCAAACCACTTGAGCTAGTGGATGGCCAAAGGTACCAAGTC	782	
Db	941	TGCCAAGATGGAGTTTACCCAAACCTCTCGAGCTAGTCCATGCCAGTGGTGCCCAAGTC	1000	
QY	783	GATAGTTCTTAGAGTGGATGAATAATCGTATGATGGGCACTCTTAAAGTACGTTAATGG	842	
Db	1001	GATAGTGTGTAGAGTGGATGAATAATCAAGCTCATGATAGCACTCTTAACTACGTTGGTGG	1060	
QY	843	AACCTGTACAGCAACTT---ACCAAAATGCCATGTTCTCTCAAGTTATATATTTCIACCTA	899	
Db	1061	GAGCTGTACAGCAACTTATAACCAAAATGCCATGTTCTCTCACTATATATGTCIACCTA	1120	
QY	900	TTATAATTAATGTCTAATGCTTGGTGATGATATTTGAAGGATATCTAA	945	
Db	1121	TTATAATTAATGCTTAACTCTTGGTGATGATATTTGAAGGATATCTGA	1166	

RESULT 3  
AAQ56672  
ID AAQ56672 standard; CONA; 1195 BP.

XX	AAQ56672;
AC	
XX	04-SEP-1994 (first entry)
XX	
DE	Sequence of Phytolacca antiviral protein (PAP) cDNA.
XX	
KW	Antiviral protein; PAP; virus-resistance; transgenic plant; ss.
XX	
OS	Phytolacca americana L.
XX	
XX	Key Location/Qualifiers
FH	33..974
FT	/*tag= a
FT	
XX	
XX	EP585554-A.
PN	
XX	
XX	09-MAR-1994.
PD	
XX	
XX	30-JUN-1993; 93EP-0110445.
XX	
PF	
XX	
XX	16-AUG-1992; 92KR-0014895.
PR	
XX	
XX	(JTNK-) JIN RO LTD.
PA	(SHIN-) SHINRO KK.
XX	
XX	Choi K, Jeon H, Jeong H, Kim M, Lee K, Moon Y;
PI	Na B;
PI	
XX	
XX	WPI; 1994-076002/10.
DR	
DR	P-PSDB; AAK48548.
XX	
XX	Expression vector for Phytolacca antiviral protein - used for
PT	producing transgenic virus-resistant plants and for producing the
PT	antiviral agent.
PT	
XX	
XX	Disclosure; Fig 1; 15pp; English.
PS	
XX	
CC	To isolate PAP gene, total cellular mRNA was purified from leaves of
CC	Phytolacca americana L. obtd. in Korea. A cDNA library was
CC	constructed. The PAP gene was selected by immunoscreening employing
CC	anti-PAP antibody. A deletion mutant was prepd. from the isolated
CC	PAP gene, and the DNA sequence of the PAP gene was determined.
XX	
SQ	Sequence 1195 BP; 420 A; 201 C; 230 G; 344 T; 0 other;

Query Match	66.8%	Score 630.8	DB 15	Length 1195		
Rest Local Similarity	79.9%	Pred. No. 5.6e-166				
Matches	756	Conservative 0	Mismatches 187	Indels 3	Gaps 1	
QY	3	GAAGGTGATGCTGTGTA	CTTGGTGAUGT	TAATACGGTGGCTCAT	TCTCGACCAACATTC	62
DB						
QY	29	GAGATGAAGTCGATGCT	TGTGGTGACAAATCA	ATATGGCTCATCTT	CGACCAACATTC	88
DB						
QY	63	AACTTGTGCGCATTAAT	ACGATCACCTTTGAT	CGCTGGAAATGCCAC	CAATATACCAATATGC	122
DB						
QY	89	AACTTGGCGCTGAATA	CAACATCATCTACA	ATGTTGGAAGTACC	ACCATTATGCAATATGCG	146
DB						
QY	123	CACCTTTATGGAATCT	CTTGGTAAATCAAG	CGAAGATCCAAAC	CTATAAATGCTATGAT	182
DB						
QY	149	CACCTTTTGTGAATCA	TCCTCGTAATGAAG	CGCAAGATCCCAAC	TTTAAATGCTATGCAAT	208
DB						
QY	183	ACCAATGCTACCTGAT	CACTAAATTCGAC	CGCCCTTAAGTACT	TTATGGTTAATGCTCAAACTGC	242
DB						
QY	209	ACCAATGCTGCCCAAT	ACAATACAAATCCA	AACTAGCTGTGTTC	GTGAGCTGCAATATTC	268
DB						
QY	243	AAAGCTAAAAACCAT	TAATACCTGAGAC	CGAAATACCTTAT	ACCGCTGATATATTC	302
DB						
QY	269	AAATAAAAAACCAT	CACATAATGCTG	ACAGGAACAAAT	TGATGCTGATGCTATATTC	328
DB						
QY	303	TGATCCCTTCAATGG	CAANTAACTGCTCT	TACCATATATTT	TAATGATATTTACAAATATGCA	362
DB						
QY	329	TGATCCCTTTCAAAC	CACTAAATGTCTT	TACCATATTT	TAATGATATTTACAAATATGCA	388
DB						

RESULT 3  
AAQ56672

ID AAQ56672 standard; CONA; 1195 BP.







AAZ59221 standard; cDNA; 1379 BP.  
AAZ59221:  
20-APR-2000 (first entry)  
Variant pokeweed antiviral protein spring leaf form coding sequence.  
Pokeweed antiviral protein spring leaf form; PAP; transgenic plant;  
resistance; potato virus X; potato virus Y; potato leaf roll virus;  
tuber; ss.  
Phytolacca americana.  
US6015940-A.  
18-JAN-2000.  
07-APR-1992: 92US-0865169.  
07-APR-1992: 92US-0865169.  
(MONS ) MONSANTO CO.  
Kaniowski WK, Tumer NE, Lodge JK:  
WPI; 2000-126326/11.  
Production of transgenic potato plants or tubers expressing pokeweed  
antiviral protein which are resistant to potato virus X or Y -  
Claim 7; Fig 5; 30pp: English.  
This is the coding sequence for a variant spring leaf form of the  
pokeweed antiviral protein (PAP) which is used to generate transgenic  
potato plants. PAP is able to confer resistance to infection by potato  
virus X (PVX) or potato virus Y (PVY) or potato leaf roll virus (PLRV)  
in the potato plant or tuber expressing PAP. PAP varies from PAP  
(AAZ59220) by mutations L20R and Y49H.  
Sequence 1379 BP; 489 A; 234 C; 269 G; 387 T; 0 other;

Query Match 66.8%; Score 629.2; DB 21; Length 1379;  
Best Local Similarity 79.8%; Pred. No. 1.7e-165;  
Matches 755; Conservative 0; Mismatches 188; Indels 3; Gaps 1;  
3 GAAGGTCATCTTCTAGTCTGCTGACCTTAATAGCGTGGCTCATCTGCTGACCACTTC 62  
221 GAAGATCAAGTCAATGCTTGTGGTGACAAATCAATATGCTGCTCATCTTGCACCACTTC 280  
63 AACTTGTGCCATAAATACGATCACCCTTTGTGCTGGAAATGCCACCATTAAACAAATATGC 122  
281 AACTTGGCTGTGAATACATCATCTACANTGTTGGAAGTACCACCATAGCAATATACGC 340  
123 CACCTTTATGAATCTCTTCTGAATCAAGGAAAGATCCAAATCAAAATCAATGCTATGCCAT 382  
341 CACTTTTTCGAATGATCTTCTGAATGAAGGAAAGATCCCAAGTTTAAATGCTATGCCAT 400  
183 ACCAATGCTACCTGATCTACTTAATTCGACCCCTTAAGTCTTATGCTTAAAGTCCAAAGTGC 242  
401 ACCAATGCTGCCCAATACAATACAAATCAAAATCAAAATGCTTGTGCTGCTGCTGCTGCTG 460  
243 AAACCTTAAACCAATACATTAATGCTGACACCAATTAATTAATGCTGCTGCTGCTGCTGCTG 302  
461 AAATAAAAAACCAATACATTAATGCTGACACCAATTAATTAATGCTGCTGCTGCTGCTGCTG 520  
303 TGATCCCTTCAATGGCAATAGTCTGCTTACCATATATTAATGATATTAATGATATTAATGAT 362  
521 TGATCCCTTCAATGGCAATAGTCTGCTTACCATATTAATGATATTAATGATATTAATGAT 580  
363 ACACCTGATGCTGACGAAATACCTTCTGCTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 422  
581 ACGCCAAGATGATACAGACTACTCTTTCGCCCAATGCCCAATTCGCTGCTTACTAAACAT 640

QY 423 TAACTACAATAGCTTATATCCGACCATGCAAAAGAAAGCAGAGTAAGAACTCAAGAAATCA 482  
DB 641 ARACTTTGATAGTGGATATCCAAACATGGAATCAAAAGCGGAGTAAGAAATCAAGAAATCA 700  
QY 483 ACHCCAAATGGGAATTCAAATACTCAGCAGTACATTTGGAAATCTCTGGACATTAATTC 542  
DB 701 GUTCCAACTGGGAATTCAAATACTCAGCAGTACATTTGGAAATCTCTGGACATTAATTC 760  
QY 543 ATTCCCTGTAAAACTGAGGCTTTTCTACTGCTAGCCTCAATGCTTCAAGG 602  
DB 761 ATTCACTGAGAAACCGAAGCGGAATTCCTATTTGGTAGCCATACAATGCTATCAGAGC 820  
QY 603 AGCGGATTCACTACATACAGAACCAAGTCAAGACTAAATTTTAATAGAGCATTTTACCC 662  
DB 821 AGCAAGATTCAAGTACATAGAGAACTAGAGTGAAGAACTAATTTTAACAGAGCATTTA 880  
QY 663 TGATCCCAAGTAAATTAATTTGGAGGAGAGTGGGCAAAATCTCTGAGCAATTAATTC 722  
DB 881 TAATCCCAAGTAAATTAATTTGGAGGAGAGTGGGCAAAATCTCTGAGCAATTAATTC 940  
QY 723 TGCCCAAGTAAATTAATTTGGAGGAGAGTGGGCAAAATCTCTGAGCAATTAATTC 782  
DB 941 TGCCCAAGTAAATTAATTTGGAGGAGAGTGGGCAAAATCTCTGAGCAATTAATTC 1000  
QY 783 CATAGTTCTTAGAGTGGATCAATCAATCGTGTGATGGGCACTTCAAGTAACTTAAATCG 842  
DB 1001 GATAGTTCTTAGAGTGGATCAATCAATCGTGTGATGGGCACTTCAAGTAACTTAAATCG 1060  
QY 843 AACTGCTGACACAACTT---ACCAAAATGCCATGTTCTCTCAAGTATTAATTTCAATTTA 899  
DB 1061 GAGCTGTCAGACAACTTATAACCAAAATGCCATGTTTCCCAACTTATAATCTCTACTTA 1120  
QY 900 TTATAATATATCTTAATCTGCTGATCTGCTGATCTGCTGATCTGCTGATCTGCTGATCT 945  
DB 1121 TTATAATATATCTTAATCTGCTGATCTGCTGATCTGCTGATCTGCTGATCTGCTGATCT 1166  
RESULT 7  
AA043967  
ID AA043967 standard; DNA; 2472 BP.  
XX AC AA043967;  
AC AC AA043967;  
DT 09-NOV-1993 (first entry)  
XX XX Pokeweed antiviral protein.  
XX KW Pokeweed; ricin; protein synthesis inhibitor; cancer;  
XX KW polymerase chain reaction; PCR; ss.  
XX OS Phytolacca americana.  
XX PH Key Location/Qualifiers  
FT CAAT\_signal 549..552  
FT CAAT\_signal /\*tag= a  
FT CAAT\_signal 627..630  
FT TATA\_signal /\*tag= b  
FT TATA\_signal 845..850  
FT sig\_peptide 1014..1085  
FT mat\_peptide 1086..1868  
FT polyA\_signal /\*tag= c  
FT polyA\_signal 2130..2135  
XX JF05137580-A.  
PN 01-JUN-1993.  
XX 20-NOV-1991; 91JP-0329672.



cells and purifying the protein from inclusion bodies. The  
antiviral proteins and recombinant proteins inhibit protein  
synthesis. They can be used as active ingredients of antiviral  
agents of plant viruses, and employed in the manufacture of  
immunoconjugates for the treatment of AIDS and cancer. The  
isolated genes can be used in the breeding of transgenic plants  
having viral resistance.

Query Match 62.3%; Score 588.6; DB 19; Length 882;  
Best Local Similarity 80.3%; Pred. No. 3e-154;  
Matches 703; Conservative 0; Mismatches 169; Indels 3; Gaps 1;

74 TAAATACGATCCTTGGTCTGGAATGCCACCAATTAACAATATGCCACCTTTATCG 133  
5 TGAATACCATCTACATCTGTTGAAGTACCACCATAGCAATAGCCACTTTTCTCG 64  
134 ATCTCTTCTGTAATCAAGCAAGATCCAAATGCTATGCTATGCAATACCAATGCTAC 193  
65 ATAATCTTCTGTAATCAAGCAAGATCCAAATGCTATGCTATGCAATACCAATGCTAC 124  
194 CTGATACCTAATCGACCCCTAAGTACTTATGTTAAGCTCCAGTGCACCACTTAAAA 253  
125 CCATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 184  
254 CCATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 313  
185 CCATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 244  
314 ATGGCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 373  
245 ATACCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 304  
374 TGGAGAACTACTCTTCTGCTCAAGTCTAGTCTGCTGCTGCTGCTGCTGCTGCTGCT 433  
305 TAGAGACTACTCTTCTGCTCAATCCCAATCTCTGCTGCTGCTGCTGCTGCTGCTGCT 364  
434 GCTTATATCCGACATGGAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 493  
365 GTCGATATCCACATTTGGAATCAAAAGCGGAGTAAATCAAGAGTCAAGTTCACCTGG 424  
494 GAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 553  
425 GAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 484  
554 AAACCTGAGGCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 613  
485 AGCTGAGCTGCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 544  
614 AGTACATAGAGCAATCAAGCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 673  
545 AGTACATAGAGCAATCAAGCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 604  
674 TAATTAATTTGAGGAGAGTGGGCAATCTCTCAGGCAATTCACATGCGCAAGATG 733  
605 TACTTAATTTGAGGAGAGTGGGCAATCTCTCAGGCAATTCACATGCGCAAGATG 664  
734 GGGCTTTACCCAAACCTTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 793  
665 GAGTTTACCCAAACCTTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 724  
794 GAGTGGATGAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 853  
725 GAGTGGATGAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 784  
854 CAACCTT---ACCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 910  
785 CAATTAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 844  
911 TCTCTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 945  
845 TGGCTAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 879

RESULT 9  
AA199557  
ID AA199557 standard; DNA: 2369 BP.  
XX  
AC AA199557;  
XX  
DT 08-JUN-1998 (first entry)  
XX  
DE Phytolacca insularis antiviral protein gpIP50 gene.  
XX  
KW Antiviral protein; gpIP50 gene; virucide; transgenic plant;  
KW virus resistance; immunoconjugate; AIDS; cancer; therapy; ss.  
OS  
XX Phytolacca insularis Nakai.

Key	Location/Qualifiers
CAAT_signal	983..986
*tag= a	
TATA_signal	1032..1035
*tag= b	
TATA_signal	1090..1093
*tag= c	
CDS	1128..2075
*tag= d	
(note= "Claim 3"	
sig_peptide	1128..1199
*tag= e	
mat_peptide	1200..2072
*tag= f	
polyA_signal	2142..2146
*tag= g	
polyA_signal	2263..2268
*tag= h	

EP808902-A2.  
XX  
XX 26-NOV-1997.  
XX  
XX 30-SEP-1996; 96EP-0307159.  
XX  
XX 22-MAY-1996; 96KR-0017404.  
XX (JINK-) JIN RO LTD.  
XX  
XX Choi J, Choi K, Choi Y, Hong E, Jin J, Kim C;  
XX Kim M, Lee J, Lee J, Moon Y, Song S, Yun Y;  
XX  
XX WPI: 1998-001788/01.  
XX P-PSDB: AAW26774.

Antiviral proteins of Phytolacca insularis Nakai and their genes  
useful in plant antiviral agents and immunoconjugates for the  
treatment of AIDS and cancer  
Claim 3: Fig 5; 26pp; English.

This polynucleotide comprises the Phytolacca insularis Nakai gene  
coding for a 35.7 kDa novel antiviral protein designated gpIP50  
(see AAW26774). The gpIP50 gene was isolated from leaf genomic DNA  
using the coding region of antiviral protein cIP as probe.  
Another gene (see AAW99556), encoding a 32.8 kDa protein (see AAW26774)  
designated gpIP2, has also been isolated from P. insularis Nakai.  
Also claimed are vectors encoding these antiviral proteins and host  
cells transformed or transfected with these vectors. E. coli  
XL1-Blue MRF' gpIP50 (MCM-1008J) host cells are claimed, as is a  
process for preparing antiviral protein by cultivating these cells  
and purifying the protein from inclusion bodies. The antiviral  
proteins and recombinant proteins inhibit protein synthesis. They  
can be used as active ingredients of antiviral agents of plant  
viruses, and employed in the manufacture of immunoconjugates for  
the treatment of AIDS and cancer. The isolated genes can be used

CC XX SQ	in the breeding of transgenic plants having viral resistance.									
	Query Match	62.1%	Score 586.4	DB 19	Length 2369					
	Best Local Similarity	78.9%	Pred. No. 1.8e-153							
	Matches 753	Conservative 0	Mismatches 106	Indels 15	Gaps 4					
OY	1	ATGAAGGTGATGCTTGTAGCTTTGGTGACGCTTAATAGCGTGGCTCATTTGCTGCACCAACT	60							
DB	1128	ATCAAGATGATGCTTGTGCTTGTGATCACAAATACAGCATGGCTAATTTTGCACCACT	1187							
OY	61	TCNACTTTGCC-----ATAAATAGCATCACCTTTGATGCTGGAATGCCACCATTAAC	114							
DB	1188	TCNACTTGGCCGCCAAGTCCAAATCCAATCACCTTCGAGTTGGAAATGGCAATTAAC	1247							
OY	115	AAATATGCCACCTTTATGGAATCTCTCTGTAATCAAGCGAAAGATCCAAACTAAAAATCC	174							
DB	1248	AGTACGCCNCTTTTATGAAATCTCTCCGTGATCAAGCGAAATCCAAATCTACAGTC	1307							
OY	175	TATGGCATAGCAATGCTAGCTGATTAATATTCGACCCCTAAGTACTTATTTGGTTAAGCT	234							
DB	1308	TATGCCATACCAATGTTGGCCAAATAGTAGTTTGACCCCAAGTACTTCTTGGTTCGCCTC	1367							
OY	235	CAAGTGTCAAACCTAAACACCATTAACATAATGCTGAGACGAAATAACTTTATAGTGTATG	294							
DB	1368	CAAGATTCAAGTTTAAACACCATCACATAATGTTGAAGCGAAATAACTTGTATGTTTATG	1427							
OY	295	GGCTATCTGATCCCTTCAATGGCAAATAAGTGTCTTACCATATATTTAATGATATTAACA	354							
DB	1428	GGTTATGCTGACACCTTAAT---AACAACTGCTGTATCATATATTTAAGGATATCTCA	1484							
OY	355	---AGCACCGACGCACTGATGTGGAAATACTCTTTGCTCAAGTCTAGTCTCTCGTGT	411							
DB	1485	ATACTACTGAACGAATGATGTGATGACTACTCTTTGTCCAAATATGAGTCTCTCGTGT	1544							
OY	412	GCAATGTCCTTAACATACAAATAGCTTATATCCGACCATTGGAAGAAAGACAGAGTAAC	471							
DB	1545	GCTAAATATTAGCTATGATACAGTTATCCAGCATTTGAAAAGAAAGTAGGACGA--	1601							
OY	472	TCAAGAATACTAGTCCAAATTTGGAAATTCAAATACTCAGCAGTCACATTTGAAAATCTCT	531							
DB	1602	TCAGAAGTAAAGTCCCACTCGGAATTCAAATACTCAACAGTATATTTGAAAAGATCTAT	1661							
OY	532	GGAGTTGATTCAATTCCTCTGTAATAACTTGAGCTTTTCTCTGTTAGGCATCCAAATG	591							
DB	1662	GGAGTGGATACAGTCAATGAGAAACCGAGGCCGAATCTTGTAGTAGCCATCCAAATG	1721							
OY	592	GTTTCAGAGCGCGGATTCAGATACATAGAGAACCAAGTCAAGACTAATTTTAATAGA	651							
DB	1722	CPACAGAGGCAACAAGATTCAGATACATAGAAAATCAGGTGAAGACTAATTTAATAGG	1781							
OY	652	GCATCTACCTTGATCCCAAAGTAATTAATTTGGAGGAGAAGTGGGCAAAATCTCTGAG	711							
DB	1782	GCATCTATCTTAATGCCAAGTACTTACTTTGGAGGAACATGGGGTAAGATTTCTACA	1841							
OY	712	GCAATTCACAATGCCAAGAAATGGGGCTTTACCCAAACCACTTCGAGTAGTGGATGCCAA	771							
DB	1842	GCAATTCATGATGCTTAAGAAATGGAGCTTTTAACCAACCTCTCGAGCTCATAAATGAAG	1901							
OY	772	GGTACCAAGTGGATAGTCTTACAGATGGATGAATCAATCGTGATGTGGCACTCTCTTAAG	831							
DB	1902	GGTACTAAGTGGATAGCTTTGAGAGTGGATGAATCAAACTCGATGTGGGACTCTCTAAC	1961							
OY	832	TAGCTTAATGGAACTGTCCAGACAACCTTACCAAAATGCCAATGTTCTCTCAAGTTAAT	891							
DB	1962	TAGCTTATGGGACCTGTCCAGACAACCTTACCAAGTGCATGTTCTCTCAACGTTAATG	2021							
OY	892	TCTACTTATTAATTAATGCTTAATCTTTGGTGATCTATTTGAAGGATTTCTAA	945							
DB	2022	TCCTACTTATTAATTAATAGTAAATCTTTGGTGATCAGTTTTCAGGGTCTCTCA	2075							

RESULT 10	
ABA96543	
ID	ABA96543 standard; DNA; 2369 BP.
XX	
AC	ABA96543;
XX	
DT	19-MAR-2002 (first entry)
XX	
DE	Phytolacca insularis antiviral protein (PIP)-encoding DNA.
XX	
KW	Antiviral protein; PIP; cloning vector; ds.
XX	
OS	Phytolacca insularis.
XX	
Key	Location/Qualifiers
CDS	1128..2075
FT	/Lag= a
FT	/product= "Phytolacca insularis antiviral protein, PIP"
FT	1128..1199
FT	/Lag= b
FT	mut_peptide
FT	1200..2072
FT	/Lag= c
FT	/product= "Mature PIP"
XX	
PN	KK98021298-A..
XX	
PD	25-JUN-1998.
XX	
PF	16-SEP-1996; 96KR-0040110.
XX	
PR	16-SEP-1996; 96KR-0040110.
XX	
PA	(JINR-) JIN RO LTD.
XX	
PI	MOON YH, Choi JN, YOON YC, Jin JH, Hong EJ, Lee JH, Chang JJ;
PT	Park YC, Choi GH, Kim CH, Song SG, Lee JS, Choi YD;
XX	
DR	WPI; 1999-299928/25.
XX	
PP	P-PSDB; AAM53033.
XX	
PT	Cloning vector using gene of Phytolacca insularis antiviral protein
PT	(PIP).
XX	
PS	Example 1; Fig 1; 9pp; Korean.
XX	
CC	The invention relates to cloning vectors comprising a gene encoding
CC	the Phytolacca insularis antiviral protein, PIP (ABA96543), or
CC	a fragment thereof. The present sequence represents DNA encoding PIP.
XX	
SQ	Sequence 2369 BP; 778 A; 373 C; 450 G; 768 T; 0 other:
	Query Match 62.1%; Score 586.4; DB 20; Length 2369;
	Best Local Similarity 78.9%; Pred. No. 1.8e-153;
	Matches 753; Conservative 0; Mismatches 106; Indels 15; Gaps 4;
OY	1 ATCAAGGCGATGCTTGTAGTTGGTGACGTTAATAGCGTGCATTCGCTGKCAACT 60
Db	1128 ATCAAGATGATGTTGCTTGTGTGATGACAATAACACCATGCCCTAATTTTGCACCACCT 1187
OY	61 TCACCTTTGCCC-----ATAAATAGCATCACCTTTGATGCTCGGAATGCCACGATTAC 114
Db	1188 TCACCTTGGSCGCGCAAGTCCAATGCAATCACCTTCGAGCTTGSAAATCGACCATTAAC 1247
OY	115 AAATATGCCACCTTTATGGAATCTCTTCGTAAATCAAGCGAAACATCCAAACATTAATGC 174
Db	.1248 AAGTAGCCACCTTTTATGAAATCTCTCCGTGATCAAGCGAAAGATCCAAATCTACAGTGC 1307
OY	175 TATGCCATACCAATGCTACCTGATCTAATATGCAGCCCTAAGTACTATTGGTTAAGCTG 234
Db	1308 TATGCCATACCAATGTTGCCAATACTACTTTTCACUCCCCCAAGTACCTGTTGTTGCGCHC 1367
OY	235 CAAGGTGCAANUCUATAAACCATTACACHAATGCTGAGACGAANTAACCTATACAGTGATG 294

06/02/2003

10

1368 CAAGATTCAAGTTTAAAGAACCATCACACTAATGTTGAAGCGAAATAACTGTATGTTATG 1427  
295 GGCTATTCTGATCCCTTCAATGGCAATAGTGTCTTACCACATATATTTAATGATATACA 354  
1428 GGTATGCTGACACCTATAAT--ACAGAGTGTGTTATCATATATTTAAGGATATCTCA 1484  
355 ---AGCAGCAAGCAGCTAGTGTGGAGAAATACTCTTTGCTCAAGTCTAGTTCGTGTT 411  
1485 AATACTACTGAAGAAATGATGTGATGACTACTCTTTGCTCCAAATATAGTCTCTGCTT 1544  
412 GCAATGTCCTAATCAATGATGATATATCCGACCATGGAAGAAAGAGAGAGAGTAAAC 471  
1545 GGTAAATAATATGATGATGAGGATGATCCAGCATTTGGAAAGAAAGATAGGACGA--- 1601  
472 TCAAGAAATCAAGTCCCAATTTGGGAATTAATCAATCTACAGCAGTGCATTTGGAAATCTCT 531  
1602 TCAAGAAATCAAGTCCCAATTTGGGAATTAATCAATCTACAGCAGTGCATTTGGAAATCTCT 1661  
532 GGAGTTGATTCATTTCCCTGTAAAGAACTGAGCGCTTTTCTTACGTGGTAGCCATCAAGT 591  
1662 GGAGTGGATACAGTCAATGAGAAACCGAGCGCAATTTCTTCTAGTAGCCATCAAGT 1721  
592 GTTTCAGAGGAGCGCGATTTCAAGTACATACAGAAACCAAGTCAAGTCAATTTTAATAGA 651  
1722 GTACAGAGGCAACAGATTCAGTACATAGAAATCAGGTGAGAGCTAATTTTAATAGG 1781  
652 GCATTCTACCTGATCCCAAGTAAATTAATTTGGAGGAAAGTGGGCAAAATCTCTGAG 711  
1782 GCATTCTATCTAATGCAAGTACTTAATTTGGAGGAAACATGGGCTAAGTCTTCTACA 1841  
712 GCATTTCACATGCGCAAGATGGGCTTTTACCAACCAACCTTGAGCTAGTGGATGCCAA 771  
1842 GCAATTCATGATGTAAGATGGAGCTTTAACCACCAACCTCTGGAGCTCATAAATGAAGAT 1901  
772 GGTACCAAGTGGATGTTCTTAGAGTGATGATCAATCAATCTGATGTGGCACTCTTAAAG 831  
1902 GGTACTAAGTGGATGTTCTTAGAGTGATGATCAATCAATCAACCTGATGTGGGACTCTTAA 1961  
832 TACGTTAATGAAGCTGTACAGCAACTTACCAAAATGCCATGTTCTCTCAAGTTATTAAT 891  
1962 TACGTTGAGGAGCTGTACAGCAACTTACCAAAATGCCATGTTCTCTCAAGTTATTAAT 2021  
892 TCTACTTATTAATATATATCTCTAATCTCTGATCTATTTGAAGGATTTCAA 945  
2022 TCTACTTATTAATATATAGTTAATCTCTGATGATGTTGAGGGTTCTGA 2075

RESULT 11  
AA064893  
ID AA064893 standard; CDNA; 918 BP.  
XX AA064893;  
AC AA064893;  
XX AA064893;  
DT 17-JAN-1995 (first entry)  
XX Antiviral protein of Phytolacca insularis Nakai.  
XX Antiviral; anti-microbial; bacteriophage; ribosome; inhibition;  
XX Inactivation; Phytolacca insularis Nakai; ss.  
XX Phytolacca insularis Nakai.  
XX Key Location/Qualifiers  
XX CDS 1..918  
XX /\*tag= a  
XX /product= Antiviral protein.  
XX AU648475-B.  
XX 21-APR-1994.  
XX 13-OCT-1993: 93AU-0048972.

28-AUG-1993: 93KR-0016938.  
(JINR-) JIN RO LTD.  
Hong-seob J, Kwan-ho I, Kyu-whan C, Man-keun K;  
Young-ho M;  
WPI: 1994-167840/21.  
P-PSDB: AAR54839.  
Nucleotide sequence from Phytolacca insularis encoding antiviral  
protein - encodes a ribosome inactivating protein which may be  
used in antimicrobial compsns.  
Claim 1: Page 13; 22pp; English.  
The antiviral protein encoded by this sequence comprises an amino  
acid sequence prevalent in ribosome inactivating proteins (see  
AAR54840). The protein may be used in anti-microbial compositions.  
Sequence 918 BP; 294 A; 177 C; 191 G; 256 T; 0 other;

Query Match 57.58; Score 543.8; DB 15; Length 918;  
Best Local Similarity 77.48; Pred. No. 141;  
Matches 714; Conservative 0; Mismatches 197; Indels 12; Gaps 4;  
Qy 8 TGATGCTTCTAGTGTGCTGACGTTAATAGCTGGCTCATTTGGTGCACCAACTTCAACTT 67  
Db 2 TGAAGTTGATGCTTGTGTCACATATCAGTATGGCTCATTTGGACCAACATCTACTT 63  
Qy 68 GTGCATAAATACGATCAGCTTTGATGCTGGAATGCCACCATTAACAAATATGCTACT 127  
Db 62 GGGCGGTGAATACCAATCATCTACCATGTTGCAAGTACCACCATTAGAAATATGCAACT 121  
Qy 128 TTATGGGAATCTCTGCTGAATACAGCGAAAGATCCAAACTAAATATGCTATGATACAA 187  
Db 122 TTGATA- ---CTTCTGCTGACGACGCAAGATCC- -AAGTTATGCTGATGCAATACAA 175  
Qy 188 TGCTACCTGATCTAATTTGCACTTCTGACCTTAACTTATTTGGTTAAGCTCCAAAGTCAAAAG 247  
Db 176 TGCTGCCCCAATTTGGATCAATCCAAATACATATTTGTTGAGCTCCAAAGTCAAAAG 235  
Qy 248 TAAAGCAATCTACATTAATCTTGAGACGAAATACCTTATACGATGCTGCTGATGCTGCT 307  
Db 236 AAGAGGATCATCACTAATGCTAAGACGAAACATTTATATGCTGCTGCTGCTGCTGCTGCT 295  
Qy 308 CTTCAATGCAATAGTGTGCTTACCATATTTAATGATATTTACAGACGCAAGCAAGCA 367  
Db 296 CTTACAA- ---CAATAGGTGCTGCTTCTTCACTTCTTAAAGGCTATCTCAGGCTGCTGCTGCT 352  
Qy 368 CTGATGTGGAGATCTCTTTGCTCAAGTCTCTAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCT 427  
Db 353 AAGATGTAGAGCTACTCTTTGCTCAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 412  
Qy 428 ACAATAGCTTATATCCGACCATCGAAAGAGAGAGAGTAACTTCAAGAAATCTCAAGTCC 487  
Db 413 ATGATAGTCCATATCCCAATCTGCAATCAAGAGAGAGAGAGTAACTTCAAGAAATCTCAAGTCC 472  
Qy 488 AATTGGCAATTTCAATATCTCAGCAGTGCATCTGCAAAATCTCTGCTGCTGCTGCTGCTGCTGCT 547  
Db 473 AACTGGCAATTTCAATATCTCAGCAGTGCATCTGCAAAATCTCTGCTGCTGCTGCTGCTGCTGCT 532  
Qy 548 CTGTAAAGACTGAGGCTTTTCTTCTACTGCTAGCCATCCAAATGGTTTCAAGAGAGAGAGAG 607  
Db 533 CTGAGACAAAGAGCTGAATTTCTACTGCTAGCCATCCAAATGGTTTCAAGAGAGAGAGAG 592  
Qy 608 GATTCAAGTACATGAGCAACCAAGTCAAGACTTAATTTTAATAGAGACTTCTAGCTGCTGCT 667  
Db 593 GATTCAAGTACATGAGCAACCAAGTCAAGACTTAATTTTAATAGAGACTTCTAGCTGCTGCT 652  
Qy 668 CCAAGTAAATTTTGGAGAGAGAGTGGGCGCAAAATCTCTGAGGCAATTTCAAGAGAGAGAG 727  
Db 653 CCAAGTAAATTTTGGAGAGAGAGTGGGCGCAAAATCTCTGAGGCAATTTCAAGAGAGAGAG 712

AA064893  
ID AA064893 standard; CDNA; 918 BP.  
XX AA064893;  
AC AA064893;  
XX AA064893;  
DT 17-JAN-1995 (first entry)  
XX Antiviral protein of Phytolacca insularis Nakai.  
XX Antiviral; anti-microbial; bacteriophage; ribosome; inhibition;  
XX Inactivation; Phytolacca insularis Nakai; ss.  
XX Phytolacca insularis Nakai.  
XX Key Location/Qualifiers  
XX CDS 1..918  
XX /\*tag= a  
XX /product= Antiviral protein.  
XX AU648475-B.  
XX 21-APR-1994.  
XX 13-OCT-1993: 93AU-0048972.

09978274

SRNT

Printed 11/17/2004

QY 728 AGAATGGGCTTTTACCACCAACCACTTGGAGTACCTGAGTATCCCAAGTACCACTGGATAG 787  
 DB 713 GGAATGGAGCTTTTACCACCACTTGGAGTACCTGAGTATCCCAAGTACCACTGGATAG 772  
 QY 788 TTCTTAGAGTGGATGAAATCAATCGTGTGGCAGCTCTTAAGTACCTTAATGGCACT 847  
 DB 773 TGTTCAGAGTGGATGAAATCAAGCTGTGTGCTCACTCTTAAGTACCTTAATGGCACT 832  
 QY 848 GTCAGACAACTT---ACCAAAATGCCATGCTCTCTCAAGTATTAATTTCTACTTATATA 904  
 DB 833 CCAGAGAACTTATAACCAAAATGCCATGCTTTCTCACTTATAATGCTCTACTTATATA 892  
 QY 905 ATTATATGCTTAATCTTGGTGAT 927  
 DB 893 ATTACATGCTTAATCTTGGTGAT 915

## RESULT 12

AAT04782  
 ID AAT04782 standard; DNA: 918 BP.

AC AAT04782;  
 DT 15-MAY-1996 (first entry)

DE DNA pJMC201 fragment encoding P. insularis antiviral protein.

KW pJMC201; Phytothacca insularis antiviral protein; PIP; Nakai;  
 KW primer: amplif; CamV 35S promoter; transgenic plant; potato;  
 KW Agrobacterium tumefaciens LBA 4404; plant cell transformation; ss.  
 OS Phytothacca insularis.

Key Location/Qualifiers  
 misc\_difference 913..915  
 /tag= a  
 /codon= seq.gat, aa:Ser

PN AU663031-B.

PD 21-SEP-1995.

PF 04-OCT-1994; 94AU-0074404.

PR 21-JUL-1994; 94KR-0017696.

PA (JINR-) JIN RO LTD.

PI Chul-Hwan K, Hong-Seob J, Kyu-Whan C, Man-Keun K;  
 PI Young-Ho M;

XX WPI; 1995-358858/47.

DR P-PSDB; AAR80106.

PT Recombinant DNA vector expressing Phytothacca insularis antiviral  
 protein - for the production of virus resistant transgenic plants

PS Example 1; Fig 1; 25pp; English.

XX This sequence represents a fragment of a recombinant DNA pJMC201 and  
 CC encodes Phytothacca insularis antiviral protein (PIP). This sequence  
 CC was isolated from a cDNA library of Phytothacca insularis Nakai using  
 CC the primer sequences given in AAT04782-84. The amplified sequence was  
 CC used, cloned with the CamV 35S promoter, in the production of a  
 CC transgenic plant, pref. potato, using Agrobacterium tumefaciens LBA  
 CC 4404 for plant cell transformation.

XX Sequence 918 BP; 294 A; 177 C; 191 G; 256 T; 0 other;

XX Query Match 57.5%; Score 543.8; DB 16; Length 918;

XX Best Local Similarity 77.4%; Pred. No. 1e-141;

XX Matches 714; Conservative 0; Mismatches 197; Indels 12; Gaps 4;

QY 8 TCATGCTTGTACTTGTGGTGCACCTTAATAGCTGGCTCATTTGCTGCACCAACTTCAACTT 67  
 DB 2 TCAAGTTGATGCTTGTGGTGCACCAATATCACTATGGCTCATTTGCTGCACCAACTTCAACTT 63  
 QY 68 GTGCCATAATACGATCAGCTTTGATGCTGGGAAATGCCACCACTTAACAATAATACCACT 127  
 DB 62 GGGCCCTCAATACCACTATCTACCACTTGGAACTACCACTTAACAATAATACCACTTCAACTT 121  
 QY 128 TTATGGAATCTCTTGTGAATCAAGCAAGATCCCAAACTTAATAATCTATGCTATATCA 187  
 DB 122 TTGGATA---CTTGTACTGAAGGCAAGATCC---AAGTTATGTGCTATGGAATCA 175  
 QY 188 TGCTACCTGATCTAATTCACCCCTTAAGTACTTATTTGCTTAAGCTCCCAAGTCCCAAGTC 247  
 DB 176 TGCTGCCCAATATTTGATCAAAATCAATATTTGCTTGAAGCTATCTCAGGTACTGAACTGG 235  
 QY 248 TAAACCACTTACACTAATGCTGAGACGAAATTAACCTTATACGTGATGGGCTATTTCTATC 307  
 DB 236 AAGAAGGCTACACTAATGCTTAAGCAAGCAAAATTTATATGCTGATGGGCTATTTCTATC 295  
 QY 308 CCTTCAATGCAATAGTCTCTTACCATATATTAATCATATTAATCATATTAACAACGACCA 367  
 DB 296 CCTACAA---CAATAGGTGCTTTCCATCTCTTTAAGGCTATCTCAGGTACTGAACTGG 352  
 QY 368 CTGATGCTGGAGAACTACTCTTCTCAAGTTCTAGTTCTCTGTTTGCATTTCACTTAATCT 427  
 DB 353 AAGATGTACAGACTACTCTTTGCCCAATGCGGATTTCTGCTGTGGTAAACATAAATCT 412  
 QY 428 ACAATAGCTTATATCGGACCATGGAAGCAAGACGAGAACTTAACCTCAAGAAATCAAGTCC 487  
 DB 413 ATGATAGTGCATATCCACATTTGCAATCAAAAGCAGGAGTAAATTTCAAGAAATCAAGTCC 472  
 QY 488 AATTGGGAATTCAAATACTCAGCAGTGCATTTGGAATAATCTCTGAGTTGATTTCACTTCC 547  
 DB 473 AACTGGAATTCGAATCTCAGCAGTGGCATTTGGAGGATTTCTGAGTGCATCTATTTCA 532  
 QY 548 CTGTAAAACTGAGGCTTTTCTTCTACTGGTGGTACCAATGCTTTCAGAGGAGGCTC 607  
 DB 533 CTGAGAGAACCGAAGCTGAATTTCTTCTACTGCTAGCCATACAAATGCTTATCAGAGTCA 592  
 QY 608 GATTCAAGTACATAGAGAACCAAGTCAAGCTAATTTTAATAGAGCATTTCACTTCACTTCC 667  
 DB 593 GATTCAAGTACATAGAGAACCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 652  
 QY 668 CCAAGTAATTAATTTGGAGGAGAGTGGGCAAAATCTCTCAGGCAATTTCAACATGGCCA 727  
 DB 653 CCNAAGTACTTATATTTGAGGAGACATGGGCTAAGATTTCTTCAACCAATTTCACTTCA 712  
 QY 728 AGAATGGGCTTTTACCCAAACCACTTTGAGTGGATGCTGAGTCCCAAGGTACCAAGTATAG 787  
 DB 713 GGAATGGAGCTTTTACCCAACTCTCTACAGCTAGTGCATGCTGCAATGGTGCATTTGCA 772  
 QY 788 TTCTTAGAGTGGATGAAATCAATCTGATGCTGGCAGTCTCTTAAGTACGTTAATGCAACT 847  
 DB 773 TGTTCAGAGTGGATGAAATCAAGCTGATGCTGCTCAAGTATTAATTTCTACTTATATA 832  
 QY 848 GTCAGACAACTT---ACCAAAATGCTGCTTCTCTCAAGTATTAATTTCTACTTATATA 904  
 DB 833 GCCAGAGAACTTATAACCAAAATGCCATGTTTCTCACTTATATGCTCTACTTATATA 892  
 QY 905 ATTATATGCTTAATCTTGGTGAT 927  
 DB 893 ATTACATGCTTAATCTTGGTGAT 915

## RESULT 13

AZ45198

ID AZ45198 standard; DNA: 934 BP.

XX AZ45198;

XX AZ45198;

DT 29-FEB-2000 (first entry)

SRRT

Printed 11/17/2004







09978274

SRNT

Printed: 11/17/2004

06/02/2003



Db 59 IPMLPNTNPKYVVLVELOSNKKTITLMLRNNLYVMGYSDPFTNKKRYHIFNDISGT 118  
Qy 121 ERTDVENTLCCSSSSRRVAMSIYNKSLYPTMKKAEVNSRNQVQLGQILSSDGIKISGVD 180  
Ub 119 ERQDVETLCPNANSRVSKINFDSPRTLESKAGVSRQVQLGQILSSDGIKISGVD 178  
Qy 181 SPVKTAEFFLLVAIQHVSAAAFKYEENOVKTNFNAFYDPKVINLEKWKIKISAIH 240  
Db 179 SFTEKTEAEFFLLVAIQHVSAAAFKYEENOVKTNFNAFYDPKVINLEKWKIKISAIH 238  
Qy 241 NAKKCALPKLELVDAKGTWKIVLRVDEINRDVALLKYVNGTCOTTY-ONAFESQVLIIST 299  
Db 239 DAKNGVLPKPLLELVDAKGTWKIVLRVDEINRDVALLKYVNGTCOTTYONAFEPQIMST 298  
Qy 300 YNYMNLGOLFEGF 314  
Db 299 YNYMNLGOLFEGF 313

RESULT 2  
S28421  
rRNA N-glycosidase (EC 3.2.2.22) PAP-alpha - Virginian pokeweed  
N:Alternate names: antiviral protein alpha-PAP  
C:Species: Phytolacca americana (Virginian pokeweed)  
C:Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 20-Jun-2000  
C:Accession: S28421  
R:Kataoka, J.; Habuka, N.; Masuta, C.; Miyano, M.; Koiwai, A.  
Plant Mol. Biol. 20, 879-886, 1992  
A:Title: Isolation and analysis of a genomic clone encoding a pokeweed antiviral protein  
A:Reference number: S28421; MUID:9309240; PMID:1281438  
A:Accession: S28421  
A:Molecule type: DNA  
A:Residues: 1-294 <RNG>  
A:Cross-references: EMBL:D10600; NID:g218010; PIDN:RAA01451.1; PID:g218011  
C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
C:Keywords: glycosidase; hydrolase  
F:30-278/Domain: rRNA N-glycosidase homology <RNG>

Query Match 68.9%; Score 1113.5; DB 2: Length 294;  
Best Local Similarity 75.4%; Pred. No. 3.9e-78;  
Matches 221; Conservative 27; Mismatches 42; Indels 3; Gaps 3;  
Qy 1 MKVMLVVVVVTLIANLIAPSTCAINTITFDAGNATINKYATFMSLRNOAKDKPKLCYC 60  
Db 3 MKVMLVVVVVTLIANLIAPSTCAINTITFDAGNATINKYATFMSLRNOAKDKPKLCYC 60  
Qy 61 IPMLPNTNPKYVVLVELOSNKKTITLMLRNNLYVMGYSDPFTNKKRYHIFNDISGT 119  
Db 61 IPMLPNTNPKYVVLVELOSNKKTITLMLRNNLYVMGYSDPFTNKKRYHIFNDISGT 119  
Qy 120 TERTDVENTLCCSSSSRRVAMSIYNKSLYPTMKKAEVNSRNQVQLGQILSSDGIKISG 179  
Db 120 TERTDVENTLCCSSSSRRVAMSIYNKSLYPTMKKAEVNSRNQVQLGQILSSDGIKISG 178  
Qy 180 DSFPVKTAEFFLLVAIQHVSAAAFKYEENOVKTNFNAFYDPKVINLEKWKIKISAIH 239  
Db 179 DSFPVKTAEFFLLVAIQHVSAAAFKYEENOVKTNFNAFYDPKVINLEKWKIKISAIH 238  
Qy 240 HNAKNGALPKLELVDAKGTWKIVLRVDEINRDVALLKYVNGTCOTTYONAF 292  
Db 239 HNAKNGALPKLELVDAKGTWKIVLRVDEINRDVALLKYVNGTCOTTYONAF 291

RESULT 3  
JE0401  
antiviral protein - Virginian pokeweed  
C:Species: Phytolacca americana (Virginian pokeweed)  
C:Date: 12-Mar-1993 #sequence\_revision 12-Mar-1993 #text\_change 07-May-1999  
C:Accession: JE0401  
R:Kung, S.S.; Kimura, M.; Funatsu, G.  
Agric. Biol. Chem. 54, 3301-3318, 1990  
A:Title: The complete amino acid sequence of antiviral protein from the seeds of pokeweed  
A:Reference number: JE0401; MUID:91242096; PMID:1368643

A:Accession: JE0401  
A:Molecule type: protein  
A:Residues: 1-261 <KUN>  
A:Experimental source: seed  
C:Comment: This protein prevents the replication of a number of plant viruses, and  
C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
F:6-254/Domain: rRNA N-glycosidase; glycoprotein  
F:10,44,255/Binding site: carbohydrate (Asn) (covalent) #status experimental  
F:34-258,84-105/Disulfide bonds: #status experimental

Query Match 67.8%; Score 1096.5; DB 2: Length 261;  
Best Local Similarity 82.8%; Pred. No. 6.7e-77;  
Matches 217; Conservative 14; Mismatches 30; Indels 1; Gaps 1;

Qy 25 INTTFDAGNATINKYATFMSLRNOAKDKPKLCYCIGIPMLPNTNPKYVVLVELOSNK 84  
Db 1 INTTFDAGNATINKYATFMSLRNOAKDKPKLCYCIGIPMLPNTNPKYVVLVELOSNK 80  
Qy 85 TITLMLRNNLYVMGYSDPFTNKKRYHIFNDISGTERTDVENTLCCSSSSRRVAMSIYN 144  
Db 61 TITLMLRNNLYVMGYSDPFTNKKRYHIFNDISGTERTDVENTLCCSSSSRRVAMSIYN 139  
Qy 145 SLPTTEKKKAEVNSRNQVQLGQILSSDGIKISGVSFPVKTAEFFLLVAIQHVSAAAF 204  
Db 120 GLYPTLEKKAGVTSRNEVQLGQILSSDGIKISGVSFPVKTAEFFLLVAIQHVSAAAF 179  
Qy 205 KYTENOVKTNFNAFYDPKVINLEKWKIKISAIHNAKNGALPKLELVDAKGTWKIVL 204  
Db 180 KYTENOVKTNFNAFYDPKVINLEKWKIKISAIHNAKNGALPKLELVDAKGTWKIVL 219  
Qy 265 RVDEINRDVALLKYVNGTCOTTT 286  
Db 240 RVDEINRDVALLKYVNGTCOTTT 261

RESULT 4  
S46239  
ribosome-inactivating proteins - Virginian pokeweed  
C:Species: Phytolacca americana (Virginian pokeweed)  
C:Date: 10-Dec-1994 #sequence\_revision 12-May-1995 #text\_change 20-May-1999  
C:Accession: S46239  
R:Boyer, J.L.; Radom, J.; Hoeveler, A.  
FEBS Lett. 347, 268-272, 1994  
A:Title: Isolation and characterization of a cDNA clone encoding the pokeweed anti-v  
A:Reference number: S46239; MUID:94307398; PMID:8034016  
A:Accession: S46239  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-310 <POY>  
A:Cross-references: GB:X78028; NID:g517180; PIDN:CAA55342.1; PID:g517181  
C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
F:29-282/Domain: rRNA N-glycosidase homology <RNG>

Query Match 31.6%; Score 511; DB 2: Length 310;  
Best Local Similarity 39.9%; Pred. No. 7.8e-32;  
Matches 130; Conservative 53; Mismatches 111; Indels 12; Gaps 14;

Qy 1 MKVMLVVVVVTLIANLIAPSTCAINTITFDAGNATINKYATFMSLRNOAKDKPKLCYC 60  
Db 3 MKVMLVVVVVTLIANLIAPSTCAINTITFDAGNATINKYATFMSLRNOAKDKPKLCYC 59  
Qy 61 IPMLPNTNPKYVVLVELOSNKKTITLMLRNNLYVMGYSDPFTNKKRYHIFNDISGT 120  
Db 60 IPMLPNTNPKYVVLVELOSNKKTITLMLRNNLYVMGYSDPFTNKKRYHIFNDISGT 113  
Qy 121 ERTDVENTLCCSSSSRRVAMSIYNKSLYPTMKKAEVNSRNQVQLGQILSSDGIKISG 178  
Db 114 SESDAQETVCPGKSKPGTQNNIPYKSKYKMSKG--GARTLGLGKIKITLKSNNK 171  
Qy 179 VDSPP-----VKTEAEFFLLVAIQHVSAAAFKYEENOVKTNFNAFYDPKVINLEK 232  
Db 172 KOATDQKQKNEAEFFLLVAIQHVSAAAFKYEENOVKTNFNAFYDPKVINLEK 231

```

Qy 233 GKISEAIHNAKG-----ALPAPLELVDAKTKWTLVRVDEINRDV-ALLKYVNGTCQTT 286
      : : | | | | : : : : : : : : : : : : : : : : : : : : : :
Db 232 DVSYSKVI--AKVYSGDSSTVTLPGLGDKDENKPKMTTATMNDLKNDI MALLTHV--TCKV- 286

Qy 287 YONAFSFOVIISTYYNMSNLGLDFE 312
      : : : : : : : : : : : : : : : : : : : : : :
Db 287 -KSSMPPEIMSYRYTSSNLGE-FE 310

RESULT 5
T12573
rRNA N-glycosidase (EC 3.2.2.22) - common ice plant
N:Alternate names: ribosome-inactivating protein
C:Species: Mesembryanthemum crystallinum (common ice plant)
C:Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 21-Jul-2000
C:Accession: T12573
R:Ripmann, J.F.; Michalowski, C.B.; Nelson, D.E.; Bohnert, H.J.
Plant Mol. Biol. 35, 701-709, 1997
A:Title: Induction of a ribosome-inactivating protein upon environmental st
A:Reference number: T17533; MUID:98087998; PMID:945592
A:Accession: T12573
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-289 <R>
A:Cross-references: EMBL:U08072; NID:g1773331; PIDN:AAB96824.1; PID:g177333
C:Genetics:
A:Gene: RIP
C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology
C:Keywords: glycosidase; hydrolase

```

```

Query Match      28.2%   Score 455.5; DB 2: Length 289;
Best Local Similarity 38.5%; Pred. No. 1.3e-27;
Matches 111; Conservative 49; Mismatches 117; Indels 11; Gaps 6;

Qy 6 VVVVTLIAMIAPTSCAINTITFDAGNATINKYATFEMSLRQADPKLKYGIPMLP 65
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 4 VLVVTLWLTMVPAAWA--TLVLDIGGATEKYSDFMTKLRTKAEAGKOPMYVGLSMMP 61

Qy 66 DTNSTPKYLLVKLOGANLKTITLMLRRNLNVYWCYSDFPNCNCKRYHIFNDITSTERTDV 125
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 62 KPTKPPYLLVLEUKVSKDMSTLSLSNDLYVVGYSOMYKG-KCRYHVFPHDHSKKPPYE 120
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 126 ETICSSSSSRVAMSYNLSLYPTMEKKAENVSRNQLQGTOLSSDIGTSGVDSPPVK 185
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 121 OHSLCDKAKDAIRKPIGVSSYTEIERKAKVYRNKEICLGYNKLTLPKYVGSESKDIQ 180

Qy 186 TEAFFLLVATOMVSEAAFEKYIENQVNTNENRAPHPPKVINLEKKWCKTSFATHNAKNG 245
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 181 DEAKFLVLAIOTTAEAAFPYIEKIAETAAN----PDQTAICLNNNSKLSKEIYNQFKG 236

Qy 246 ALPKPLEVDKAGTKVLVLVDDEINRQVALIKVYNGTC--OTTYONAM 291

Db 237 NPQTPAKDVTVRCK-RLKGINGVOKYITLLSY-OGTSGROOLTFRKVV 282

```

RESULT 6  
JC4811  
betavulgin - beet  
C: Species: Beta vulgaris (beet)  
C: Date: 15-Aug-1996 #sequence\_revision 15-Oct-1996 #text\_change 20-Jun-2000  
C: Accession: JC4811  
R: Hornung, E.; Wajant, H.; Jeske, H.; Mundry, K.W.  
Gene 170, 233-236, 1996  
A: Title: Cloning of a cDNA encoding a new ribosome-inactivating protein from Beta vulgaris  
A: Reference number: JC4811; MUID:96235141; PMID:8666251  
A: Accession: JC4811  
A: Molecule type: mRNA  
A: Residues: 1-272 <HOR>  
A: Cross-references: EMBL: X85967; NID: g1405840; PIDN: CAA5952.1; PID: g1405841  
A: Note: the source is designated as Beta vulgaris vulgaris (mangold)  
C: Comment: This protein belongs to ribosome-inactivating protein type II.  
C: Superfamily: rRNA N-glycosidase; rRNA N-glycosidase; rRNA N-glycosidase homology

F;28-271/Domain: rRNA N-glycosidase homology <RNG>

```

06/02/20
Query Match      25.1%   Score 406.   DR 2:   Length 272:
Best Local Similarity 38.8%   Pred. No. 7.7e-24:
Matches 111: Conservative 40: Mismatches 11: Indels 22: Gaps 8:

QY      1  MKYLVVVVTLIAWLAAPSTSTCAINTITFDAGNATINKYATATEMESLRNKAQDKLKCYG 60
DB      1  MKAHWLVIATSIIMVVLQGVYAA--DVTDFEYASKTKYGTFLSLNRKIVKDSKILVYRK 58

QY     61  IPMIIPDNTSPKYILVVKLQ-----CANLKTTITLMILRRNNLLYVMGYSDPFENGNCKRYIIFMD 116
DB     59  IPMLPADIPKPKAKYLLAEALKAKKAGTDI--TITLAVSKNDLYVVVAFDQVAG-KLKAHYFPH 116

QY    117  ITSTERTDVENTICSSSSSRVAMISINTNSIYPTMEKKAEVNSRNQVQIGTIGTISSDICKI 176
DB    117  ISLATAKAIEPT-----ATQYIQIGYTSNVYSIEGAAGSN--RVNFOLFGEVKLKEYMLYN 169

QY    177  SG---VDSFPVKTEAFELLVATQNVSAARKEYTENOVKTFNFRAFYDPDPKVINILEKRWG 233
DB    170  YGKNVQSDYDKRSARFELLAALQWVAARKEYVESKAINN-----VVPDYKVVSYLENHNS 225

QY    234  KISEAIHNAKNGALPKPIELVQAGKTAKWILVLRDFINRDVALLKYV 279
DB    226  KISECTRKAVKVKVISPPFELVNASNGRWVYNQVSDILKPDWGIILSYV 271

RESULT 7
S05205
rRNA N-glycosylase (BC 3.2.2.22) 6 precursor - common soapwort (fragment)
N:Alternate names: ribosome-inactivating protein SO-6; saporin 6
C:Species: Saponaria officinalis (common soapwort)
C:Date: 30-Sep-1991 #sequence,revision 30-Sep-1991 #text_change 20-Aug-1999
C:Accession: S05205; A32331; S17689; S16330; A45499
R:Benatti, L.; Saccardo, M.B.; Dani, M.; Nitti, G.; Sassano, M.; Lorenzetti, R.; Loi
Eur. J. Biochem. 183, 465-470, 1989
A:Title: Nucleotide sequence of cDNA coding for saporin-6, a type-I ribosome-inactiva
A:Reference number: S05205; MWID:09348421; PMID:2547612

```

F:30-273/Domain: rRNA N-glycosidase homology <RNG>

Query Match 23.4% Score 378.5 DB 2 Length 283;  
Best Local Similarity 36.4% Pred. No. 1e-21;  
Matches 104; Conservative 56; Mismatches 103; Indels 23; Gaps 10;

Qy 4 MLVVVTLIAWLI---AAPTSTCAINTITDAGNATINKYATFMSLRNQAOKPKLKYC 60  
Db 1 MKIYVATIAWILLOFSAWTTDAVTSITLDLVNPTAGYSSFYVDKIRNNVKNOPNLYKG 60  
Qy 61 IPM-LPOTNSTPKYLLVLOGANLTKITLMLRRNNLYVMGYSDPFNGKCR-YHIFNDIT 118  
Db 61 TDIAVIGPPSKKFLRFNFQSSR-GTVSLGLKRONLYVAVYLAAMDNTNVRAYYFKSEIT 119  
Qy 119 STERTDVENTLCSSSSRVAMSIYNLSLYPTMEKAEV---NSRNOVOLGICQILSSDIG 174  
Db 120 SAEST-----ALFPEATANOKALEYTEDYOSIEKNAQITOGDSRKELGICIDLLSTSME 175  
Qy 175 KISGVDSFPVKTEAFFLLVAIQMVSEAAKRYENOVKTNFNRAFYDDPKVINLEEKWK 234  
Db 176 AVN-KKARVVKDEARFLLIAIQMTAEAFRYIQNLVKNPKNFSENKVIQFEVANKK 234  
Qy 235 ISEAIH-NAKNGALPKPLELVDAKTKWILRVDEINRDVALLKYV-279  
Db 235 ISTAIYGDARKGVNKDYDFGFK-----VRQVKDL--QMGLLMYL 273

## RESULT 8

RLOH2

rRNA N-glycosidase (EC 3.2.2.22) Sap2 precursor - common soapwort  
N:Alternate names: ribosome-inactivating protein; saporin 2; saporin S5  
C:Species: Saponaria officinalis (common soapwort)  
C:Date: 31-Mar-1993 #sequence; revision 31-Mar-1993 #text\_change 16-Jul-1999  
C:Accession: S17933; S28540; S38526; S15459  
R:Fordham-Skelton, A. P.; Taylor, P. N.; Hartley, M. R.; Croy, R. R. D.  
Mol. Gen. Genet. 229, 460-466, 1991  
A:Title: Characterisation of saporin genes: in vitro expression and ribosome inactivating  
A:Reference number: S17932; MUID:92049247; PMID:1719367  
A:Accession: S17933

A:Molecule type: DNA  
A:Residues: 1-292 <FOR>  
A:Cross-references: EMBL:X59255; GB:S63902; NID:g2094853; PIDN:CAA41948.1; PID:g21321  
R:Soria, M. R.  
Submitted to the EMBL Data Library, October 1992  
A:Reference number: S28540  
A:Accession: S28540  
A:Molecule type: DNA

A:Residues: 25-283 <SOR>  
A:Cross-references: EMBL:X69132; NID:g21330; PIDN:CAA48886.1; PID:g938284  
R:Ferreras, J.M.; Barbieri, L.; Girbes, T.; Battelli, M.G.; Rojo, M.A.; Arias, F.J.; Rojo,  
Blochim. Biophys. Acta 1216, 31-42, 1993  
A:Title: Distribution and properties of major ribosome-inactivating proteins (28 S rRNA  
A:Reference number: S38521; MUID:94032486; PMID:8218413

A:Accession: S38526  
A:Molecule type: protein  
A:Residues: 25-54 <PER>  
C:Genetics:

A:Gene: SAP2  
C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
C:Keywords: glycosidase; hydrolase; monomer; toxin  
F:1-24/Domain: signal sequence #status predicted <SIG>  
F:25-292/Product: rRNA N-glycosidase Sap2 #status predicted <MAT>  
F:30-273/Domain: rRNA N-glycosidase homology <RNG>

Query Match 23.0% Score 372; DB 1; Length 292;  
Best Local Similarity 35.4% Pred. No. 3.4e-21;  
Matches 108; Conservative 52; Mismatches 107; Indels 36; Gaps 12;

Qy 4 MLVVVTLIAWLI---AAPTSTCAINTITDAGNATINKYATFMSLRNQAOKPKLKYC 60  
Db 1 MKIYVATIAWILLOFSAWTTDAVTSITLDLVNPTAGYSSFYVDKIRNNVKNOPNLYKG 60  
Qy 61 IPM-LPOTNSTPKYLLVLOGANLTKITLMLRRNNLYVMGYSDPFNGKCR-YHIFNDIT 118

Db 61 TDIAVIGPPSKKFLRFNFQSSR-GTVSLGLKRONLYVAVYLAAMDNTNVRAYYFKSEIT 119  
Qy 119 STERTDVENTLCSSSSRVAMSIYNLSLYPTMEKAEV---NSRNOVOLGICQILSSDIG 171  
Db 120 SAEST-----ALFPEATANOKALEYTEDYOSIEKNAQITOGDSRKELGICIDLLSTSME 175  
Qy 172 DICKISGVDSFPVKTEAFFLLVAIQMVSEAAKRYENOVKTNFNRAFYDDPKVINLEEK 231  
Db 176 AVNKKARV---VKNEARFLLIAIQMTAEAFRYIQNLVKNPKNFSDNSKVIQFEVS 231  
Qy 232 WKISRAIH-NAKNGALPKPLELVDAKTKWILRVDEINRDVALLKYV-----NC 281  
Db 232 WRKISTAIYGDARKGVNKDYDFGFK-----VRQVKDL--QMGLLMYLCKPKSSNKANS 284  
Qy 282 TCQTT 286  
Db 285 TAYAT 289

## RESULT 9

S17519

rRNA N-glycosidase (EC 3.2.2.22) dianthin 30 precursor - clove pink  
N:Alternate names: DAP30 protein; ribosome-inactivating protein dianthin 30  
C:Species: Dianthus caryophyllus (clove pink)  
C:Date: 20-Feb-1995 #sequence; revision 20-Feb-1995 #text\_change 33-Sep-1998  
C:Accession: S17519; S17685  
R:Legname, G.; Bellosta, P.; Gromo, G.; Modena, D.; Keen, J. N.; Kokoris, L. M.; Lord,  
Biochim. Biophys. Acta 1090, 119-122, 1991  
A:Title: Nucleotide sequence of cDNA coding for dianthin 30, a ribosome inactivating  
A:Reference number: S17519; MUID:91355219; PMID:1840496  
A:Accession: S17519

A:Molecule type: mRNA  
A:Residues: 1-293 <LEG>  
A:Cross-references: EMBL:X59260  
A:Experimental source: Leaf  
R:Lee-Huang, S.; Kung, H.; Huang, P. L.; Huang, P. L.; Li, B. Q.; Huang, H. J.  
FEBS Lett. 291, 139-144, 1991  
A:Title: A new class of anti-HIV agents: GAP31, DAPS 30 and 32.  
A:Reference number: S17574; MUID:92037998; PMID:1936243  
A:Accession: S17685

A:Molecule type: protein  
A:Residues: 24-40, X, 42-82 <DAP>  
C:Genetics:  
A:Gene: dia 30  
C:Function:

A:Description: removes a single adenine residue from a highly conserved loop structure  
A:Note: high antiviral potency, but low toxicity to cells in culture or intact animals  
C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
C:Keywords: antiviral; glycosidase; hydrolase; monomer; toxin  
F:1-23/Domain: signal sequence #status predicted <SIG>  
F:24-293/Product: rRNA N-glycosidase dianthin 30 #status experimental <MAT>  
F:29-273/Domain: rRNA N-glycosidase homology <RNG>

Query Match 21.4% Score 345.5; DB 2; Length 293;  
Best Local Similarity 34.6% Pred. No. 3.7e-19;  
Matches 100; Conservative 52; Mismatches 108; Indels 29; Gaps 12;

Qy 4 MLVVVTLIAWLIAPTS---TCAINTITDAGNATINKYATFMSLRNQAOKPKLKYC 60  
Db 1 MKIYVATIAWILFQSSWTTDAATATLNLNAPSASQYSSFLDQIRNNVDTSLIYQG 59  
Qy 61 --IPMLPDTNSTPKYLLVLOGANLTKITLMLRRNNLYVMGYSDPFNGKCR-YHIFNDI 117  
Db 60 TDVAVIGAPSTTDKFLRNFUGPR-GTVSLGLKRNLYVAVYLAAMDNTNVRAYYFKNO 118  
Qy 118 TSTERTDVENTLCSSSSRVAMSIYNLSLYPTMEKAEV---NSRNOVOLGICQILSSDIG 173  
Db 119 TSAETALFPEVVVAVAKO-----LEYGEDYQALTEKNAKITTDGDSRKELGICIDLLSTSME 171  
Qy 174 GKISGVDS--FVVKTEAFFLLVAIQMVSEAAKRYENOVKTNFNRAFYDDPKVINLEEK 231  
Db 172 TMIDGVNKKVYVYKDEAKFLLIAQMTAEAFRYIQNLVKNPKNFSDNSKVIQFEVS 231





C:Keywords: glycosidase; hydrolase; toxin  
F:6-249/Domain: rRNA N-glycosidase homology <RNG>

Query Match 20.8% Score 336; DB 2; Length 253;  
Best Local Similarity 35.5%; Pred. No. 1.7e-18;  
Matches 93; Conservative 51; Mismatches 98; Indels 20; Gaps 9;

QY 25 INTITDAGNATINKYATFMSLRNOAKDPKLCYGYPM-LPDTNSTPKYLLVKOGANL 83  
DB 1 VTSITDLVNPTAGQYSSFDKIRNNKDPNLKYGGTDIAVGPSPKFKFLRINFQSSK 59

QY 84 KTTTLMRLNNLYVMGYSDPFNGKCR-YHIFNDITSTERTDVENTLCSSSSSRVAMSN 142  
DB 60 GTVSLGLKRDNLVYVAYLAMDNTNVRAYFRSEITSAELT---ALFPEATNOKALE 115

QY 143 YNSLYPTMEKAEV-----NSRNOVOLGIQLSSDIGNKISGVDFPVKTEAFFLLVAIQMV 198  
DB 116 YTEDYQSIENNAQITQDCKSKKELGLDILLTSMEAVN-KKARVVKNKARFLLIATQMT 174

QY 199 SEARFYKIENOVKTNFRAFYDPKVINLEEKWKGISSEATH-NAKNGALPKPLELYDAK 257  
DB 175 AEARFYQIQLVTKNPNFNSKNKVIQFENVNKKISTALYGOAKNGVKNKYDIPGEGK 234

QY 258 GTRKIVLRVDEINRDVALLKYV 279  
DB 235 -----VRQVKDL--QMGLLHYL 249

## RESULT 14

S28541  
rRNA N-glycosidase (EC 3.2.2.22) (clone G-4) - common soapwort (fragment)  
N:Alternate names: ribosome-inactivating protein; saporin  
C:Species: Saponaria officinalis (common soapwort)  
C:Date: 07-May-1993 #sequence\_revision 07-May-1993 #text\_change 20-Aug-1999  
C:Accession: S28541

R:Sorta, M.R.  
submitted to the EMBL Data Library, October 1992

A:Reference number: S28539  
A:Accession: S28541  
A:Molecule type: DNA  
A:Residues: 1-253 <SOR>  
A:Cross-references: EMBL:X69133; NID:q21331; PIDN:CA44887.1; PID:q21332  
C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
C:Keywords: glycosidase; hydrolase; toxin  
F:6-249/Domain: rRNA N-glycosidase homology <RNG>

Query Match 20.7% Score 335; DB 2; Length 253;  
Best Local Similarity 35.8%; Pred. No. 2e-18;  
Matches 95; Conservative 47; Mismatches 97; Indels 26; Gaps 10;

QY 25 INTITDAGNATINKYATFMSLRNOAKDPKLCYGYPM-LPDTNSTPKYLLVKOGANL 83  
DB 1 VTSITDLVNPTAGQYSSFDKIRNNKDPNLKYGGTDIAVGPSPKFKFLRINFQSSR 59

QY 84 KTTTLMRLNNLYVMGYSDPFNGKCR-YHIFNDITSTERTDVENTLCSSSSSRVAMSN 142  
DB 60 GTVSLGLKRDNLVYVAYLAMDNTNVRAYFRSEITSAELT---ALFPEATNOKALE 115

QY 143 YNSLYPTMEKAEV-----NSRNOVOLGIQL---SSDIGNKISGVDFPVKTEAFFLLVAI 195  
DB 116 YTEDYQSIENNAQITQDCKSKKELGLDILLTFMEAVNKKARV---VKNAREFLLIAT 171

QY 196 QMVSEARFYKIENOVKTNFRAFYDPKVINLEEKWKGISSEATH-NAKNGALPKPLELY 254  
DB 172 QMTAEARFYQIQLVTKNPNFNSKNKVIQFENVNKKISTALYGOAKNGVKNKYDIPG 231

QY 255 DAKGTRKIVLRVDEINRDVALLKYV 279  
DB 232 FGK-----VRQVKDL--QMGLLHYL 249

## RESULT 15

JT0753

rRNA N-glycosidase (EC 3.2.2.22) precursor - Gelonium multiflorum  
N:Alternate names: gelonin; type I ribosome-inactivating protein  
C:Species: Gelonium multiflorum  
C:Date: 30-Sep-1993 #sequence\_revision 20-Aug-1994 #text\_change 20-Aug-1999  
C:Accession: JT0753; S16489  
R:Nolan, P.A.; Garrison, D.A.; Better, M.  
Gene 134, 223-227, 1993

A:Title: Cloning and expression of a gene encoding gelonin, a ribosome-inactivating  
A:Reference number: JT0753; MUID:94085781; PMID:7916721

A:Molecule type: mRNA  
A:Accession: JT0753  
A:Residues: 1-316 <NDL>  
A:Cross-references: GB:L12243; NID:q388633; PIDN:AAA16312.1; PID:q388634  
R:Montecuchi, P.C.; Lezzarini, A.M.; Barberi, L.; Stirpe, F.; Sorio, M.; Lippi,  
Int. J. Pept. Protein Res. 33, 263-267, 1989  
A:Title: N-terminal sequence of some ribosome-inactivating proteins.  
A:Reference number: S16331; MUID:89326691; PMID:2753596  
A:Accession: S16489  
A:Molecule type: protein  
A:Residues: 47-89, 'K', 91-92, 'D' <MON>  
C:Function:

A:Description: hydrolyzes the N-glycosidic bond of a specific adenosine in 28S rRN  
C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
C:Keywords: glycosidase; hydrolase  
F:1-46/Domain: signal sequence #status predicted <SIG>  
F:47-316/Product: ribosomal RNA N-glycosidase #status predicted <MAT>  
F:53-294/Domain: rRNA N-glycosidase homology <RNG>

Query Match 20.4% Score 330; DB 2; Length 316;  
Best Local Similarity 30.8%; Pred. No. 6.4e-18;  
Matches 97; Conservative 60; Mismatches 122; Indels 36; Gaps 11;

QY 7 VVTLIAWLLAAPT-----TCAINITFDACNATINKYATFMSLRNOAKDPKLCY 59  
DB 23 IVIGSTARIFSLPTNDEETSRTGLDVTFSFTKGATYJTYVNFLEIRVKLK-PEGNSH 81

QY 60 GIPMLPDTNSTPK--YLLVKLOGANLKTITLMRLNNLYVMGYSDPFNGKCRHYIENDI 117  
DB 82 GIPLLKKKCDPCPKCFVLVALSNDNGOLAEITADVTSVYVGYO-----VNRKSYFFKDA 136

QY 118 TSTERTDV-ENTLCSSSSSRVAMSNINSLYPTMEKAEVNSRNOVOLGIQLSSDICKT 176  
DB 137 PDAAEGLFKNT-----IKTRLHFGGSYPSLE--GEKAYRETTDLGTEPLRIGIKKL 186

QY 177 --SGVDSFPVKTEAFFLLVATQMVSEARFYKIENOVKTNFRAFYDPKVINLEEKWK 234  
DB 187 DENAIQNKPTETIASLLVVIOMVSEARFTIFENIRNNFOQIRPANNTISLENKWK 216

QY 235 ISEAIH-NAKNGALPKPLELYDAKTKWIVLRVDEINROVALLKYVNGTCQTTYQAMNS 293  
DB 247 LSFQIRTSGANCMFSEAVELERANGKYYVFAVDQVKPKIALLKFDVKDKPT----SLAA 302

QY 294 QVLIISTYNYMNLG 308  
DB 303 ELTIQ-----NYESLVG 314

Search completed: May 28, 2003, 09:59:04  
Job time : 47 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using .sw model

Run on: June 2, 2003, 02:32:48 ; Search time 2636 Seconds  
(without alignments)  
10433.290 Million cells

US-09-978-274A-1

Title: 05-0  
 Perfect score: 945

Sequence: 1 atqaaggtgattcttgaat.....atctatttqaaggaattctaa 945

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Minimum	DB seq	length:
Maximum	DB seq	length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl : \*

pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
C	1	945	100.0	945	6	AX427702	Sequence
	2	841	89.0	1249	8	PAPASRP1	X98079 P. americana
	3	737	78.0	792	6	AX427704	Sequence
	4	712	75.3	1092	6	AX427720	Sequence
	5	494	52.3	780	8	AB071855	Phytolacc
	6	459	48.6	465	6	AX427706	Sequence
	7	329	34.8	333	6	AX427708	Sequence
	8	42	4.4	42	6	AX427716	Sequence
	9	42	4.4	42	6	AX427717	Sequence
	10	42	4.4	783	8	AB071854	Phytolacc
C	11	32	3.4	2472	6	E05033	AD5033 DNA encodin
	12	32	3.4	2472	6	PICAPAP	D10600 P. american
	13	32	3.2	52	6	AX427714	Sequence
	14	30	3.2	1164	8	AY049785	Phytolacc
	15	29	3.1	1164	8	PAPAP	X55183 P. americana
	16	29	3.1	1195	6	A36639	Sequence 1
	17	29	3.1	1195	6	A42103	Sequence 1
	18	29	3.1	1195	6	I43835	Sequence 1
	19	29	3.1	1195	6	I58866	Sequence 1
	20	29	3.1	1378	6	AX427732	Sequence
C	21	29	3.1	1379	6	AR009535	Sequence
	22	29	3.1	1379	6	AB136704	Sequence
	23	29	3.1	1379	6	AB136705	Sequence
	24	29	3.1	1379	6	AR141172	Sequence
	25	29	3.1	1379	6	AX427731	Sequence
	26	27	2.9	47	6	AX427713	Sequence
	27	27	2.9	49	6	AX427715	Sequence
	28	26	2.8	48	6	AX427712	Sequence
	29	25	2.6	45	6	AX427721	Sequence
	30	24	2.5	43	6	AX427710	Sequence
C	31	24	2.5	43	6	AX427711	Sequence
	32	24	2.5	1387	8	AF445416	AF445416 Bombyx m
	33	24	2.5	882	6	AF67183	AF67183 Sequence 1
	34	24	2.4	34578	9	AC107069	AC107069 Homo sapi
	35	23	2.4	46778	2	AC079784	AC079784 Homo sapi
	36	23	2.4	180771	2	AC099611	AC099611 Mus muscu
	37	22	2.3	918	6	A43003	A43003 Sequence 1
	38	22	2.3	918	6	A48150	A48150 Sequence 1
	39	22	2.3	918	6	I60482	I60482 Sequence 1
	40	22	2.3	918	6	I89987	I89987 Sequence 1
C	41	22	2.3	951	6	A67185	Sequence 3
	42	22	2.3	2369	8	AF141331	AF141331 Phytolacc
	43	22	2.3	154737	2	AC129370	AC129370 Rattus no
	44	21	2.2	171568	2	AC102477	AC102477 Mus muscu
	45	21	2.2	188547	2	AL727395	AL727395 Mus muscu

## ALIGNMENTS

RESULT 1  
AX427702

AX427702	AX427702	945 bp	DNA	linear	PAJ 20-JUN-2002
LOCUS	Sequence 1 from Patent WO0233107				
DEFINITION	AX427702				
ACCESSION	AX427702				
VERSION	AX427702.1	GI:21537815			
KEYWORDS					
SOURCE	Virginian pokeweed				
ORGANISM	Phytolacca americana				
	Eukaryota: Viridiplantae; Streptophyta; Tracheophyta;				
	Angiosperms: Malvales: Malvaceae; Phytolacaceae;				
	Phytolacaceae; Phytolacaceae; Phytolacaceae;				
REFERENCE	1				
AUTHORS	Neelam, A., Atkinson, H. J., McPherson, M. J., and Thomas, C. J. R.				
TITLE	Plant cell death system				

JOURNAL Patent: WO 0233107-A 1 25-APR-2002;  
CAMBRIDGE ADVANCED TECH (GB)  
FEATURES Location/Qualifiers  
source 1..945  
/organism="Phytolacca americana"  
/db\_xref="taxon:3527"  
misc\_feature 1..24  
/note="Binding site for primer PPS18P"  
misc\_feature complement(735..776)  
/note="Binding site for primer PSXDR"  
misc\_feature 736..777  
/note="Binding site for primer PSXDF"  
variation 750..759  
/note="Sequence replacing removed XbaI site"  
misc\_feature complement(922..945)  
/note="Binding site for primer PPS2SR"  
BASE COUNT 307 a 187 c 180 g 271 t  
ORIGIN  
Query Match 100.0%; Score 945; DB 6; Length 945;  
Best Local Similarity 100.0%; Pred No. 0;  
Matches 945; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ATGAAGTGATGCTGTGAGTGTGCTGACGTTAATAGCGTGCCTCATTTGCTGCCAACCT 60  
Db 1 ATGAAGTGATGCTGTGAGTGTGCTGACGTTAATAGCGTGCCTCATTTGCTGCCAACCT 60  
Qy 61 TCACTGTGCGCAATTAATACGATCACCTTTGATGCTGGAATGCCACCATTAACAAATAT 120  
Db 61 TCACTGTGCGCAATTAATACGATCACCTTTGATGCTGGAATGCCACCATTAACAAATAT 120  
Qy 121 GCCACCTTTATGGAATCTCTCGTAATCAACGGAAGATCCAAACTAAAATGCTATGGC 180  
Db 121 GCCACCTTTATGGAATCTCTCGTAATCAACGGAAGATCCAAACTAAAATGCTATGGC 180  
Qy 181 ATACAATGCTACCTGATGACTAATTCGACCCCTTAAGTACTTATTGGTTAAGCTCAAGGT 240  
Db 181 ATACAATGCTACCTGATGACTAATTCGACCCCTTAAGTACTTATTGGTTAAGCTCAAGGT 240  
Qy 241 GCAACCTTAAACCAATTAATGCTGAGAGCAATTAATGCTGATGCTATGCTGCTATGCTAT 300  
Db 241 GCAACCTTAAACCAATTAATGCTGAGAGCAATTAATGCTGATGCTATGCTGCTATGCTAT 300  
Qy 301 TCTGATCCTTCAATGCGCAATGCTGCTTACCATAATTAATGATATTTACAGCACC 360  
Db 301 TCTGATCCTTCAATGCGCAATGCTGCTTACCATAATTAATGATATTTACAGCACC 360  
Qy 361 GAACCCACTGATGGGAGAACTACTCTTTGCTCAAGTCTAGTCTGCTGTTGCTGCTATGCTC 420  
Db 361 GAACCCACTGATGGGAGAACTACTCTTTGCTCAAGTCTAGTCTGCTGTTGCTGCTATGCTC 420  
Qy 421 ATTAACATAATAGCTTATATCCGACCATGGAAGGAGAGGAGAGTAACTCAAGAAAT 480  
Db 421 ATTAACATAATAGCTTATATCCGACCATGGAAGGAGAGGAGAGTAACTCAAGAAAT 480  
Qy 481 CAAGTCCAAATGGGAATTCAAATACTACGAGTGCATTTGGAATAATCTCTGGAGTTGAT 540  
Db 481 CAAGTCCAAATGGGAATTCAAATACTACGAGTGCATTTGGAATAATCTCTGGAGTTGAT 540  
Qy 541 TCATTCCCTGTAACCACTGAGGCTTTTCTCTACTGCTGAGCCTCAATCCCAATGTTTCAGAG 600  
Db 541 TCATTCCCTGTAACCACTGAGGCTTTTCTCTACTGCTGAGCCTCAATCCCAATGTTTCAGAG 600  
Qy 601 CGACGGCATTAAGTACATAGAGAACCAAGTCAAGACTAATTTTATAGAGCATTTCTAC 660  
Db 601 CGACGGCATTAAGTACATAGAGAACCAAGTCAAGACTAATTTTATAGAGCATTTCTAC 660  
Qy 661 CCTGATCCCAAGTAATTAATTTGGAGAGAGTGGGCGCAAAATCTCTGAGGCAATTTCCAC 720  
Db 661 CCTGATCCCAAGTAATTAATTTGGAGAGAGTGGGCGCAAAATCTCTGAGGCAATTTCCAC 720  
Qy 721 AATGCCAAGATGGGCTTTTACCCAAACCACTTACCTAGTGTAGTGGATGCCAAAGTACCAAG 780  
Db 721 AATGCCAAGATGGGCTTTTACCCAAACCACTTACCTAGTGTAGTGGATGCCAAAGTACCAAG 780

Db 721 AATGCCAAGATGGGCTTTTACCCAAACCACTTACCTAGTGTAGTGGATGCCAAAGTACCAAG 780  
Qy 781 TCGTACTTCTTAGAGTGGATGAAATCAATCGTATGTCGACTCCTTACCTTACCTTAAAT 840  
Db 781 TCGTACTTCTTAGAGTGGATGAAATCAATCGTATGTCGACTCCTTACCTTAAAT 840  
Qy 841 GGAACCTGTGACACACTTACCAAAATGCCATGCTCTCAAGTATATAATTTTACTTAT 900  
Db 841 GGAACCTGTGACACACTTACCAAAATGCCATGCTCTCAAGTATATAATTTTACTTAT 900  
Qy 901 TATAATATATGCTAATCTTGTGATCTATTTGAAGGATCTAA 945  
Db 901 TATAATATATGCTAATCTTGTGATCTATTTGAAGGATCTAA 945  
RESULT 2  
PAPASRIP 1249 bp mRNA linear PLN 30-APR-1997  
LOCUS P.americana mRNA for pokeweed antiviral protein.  
DEFINITION X98079  
ACCESSION X98079  
VERSION X98079.1 GI:1707648  
KEYWORDS PAP-S gene; pokeweed antiviral protein (PAP); ribosome-inactivating protein.  
SOURCE Phytolacca americana  
ORGANISM Phytolacca americana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Caryophyllidae; Caryophyllales; Phytolaccaceae; Phytolacca.  
REFERENCE 1 (bases 1 to 1249)  
AUTHORS Poyet,J.L. and Hoeweler,A.  
TITLE cDNA cloning and expression of pokeweed antiviral protein from seeds in Escherichia coli and its inhibition of protein synthesis in vitro  
JOURNAL FEBS Lett. 406 (1-2), 97-100 (1997)  
MEDLINE 97263479  
PUBMED 9109394  
REFERENCE 2 (bases 1 to 1249)  
AUTHORS Poyet,J.L.  
TITLE Direct Submission  
JOURNAL Submitted (22-MAY-1996) J. L. Poyet, Laboratoire de Biochimie-Biologie Molculaire, UFR Sciences et Techniques, 16 route de Gray, 25030 Resancon Cedex, FRANCE  
REMARK Revised by author 20-SEP-1996  
FEATURES Location/Qualifiers  
1..1249  
/organism="Phytolacca americana"  
/db\_xref="taxon:3527"  
/tissue\_type="seeds"  
106..1050  
/gene="PAP-S"  
106..1050  
/gene="PAP-S"  
/note="ribosome-inactivating protein type 1"  
/codon\_start=1  
/product="pokeweed antiviral protein"  
/protein\_id="CA66702.1"  
/db\_xref="GI:1707649"  
/db\_xref="SPTR:EMBL:P93444"  
/translation="MKVMI.VVVVT.LIAHLIAAPTSTCAINTIDFDAGNATINKYATFEM  
ESLRNQAQPKLKYCIGIPMLPDYNSTPKYLLVKGANLKTITLMLKNNLYVMCYSD  
PFENKCKYHIFNDITSTERTDENTLCSSSRVAMNSINSLYPTTMEKKAENSRN  
QVQIGIOJLSSDIOKISGVDSFPVKTEAFFLLVNIOMVSEARFKYTENGJVKTNINHA  
FYDPDKVINLEEDKIGKISEAHKNAKNGALPKPLELVDKAGTKNIVIRVIDEINRUVAL  
KYVNGTCQITTYQNAFTSQVITSTYYNYSNLSGIDLEFGP"  
106..177  
/gene="PAP-S"  
178..1047  
/gene="PAP-S"  
/product="unnamed"  
1212..1217  
polyA\_site  
1235  
BASE COUNT 418 a 236 c 241 g 354 t  
ORIGIN

Query Match 89.0%; Score 841; DB 8; Length 1249;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 941; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 ATCAAGGTGATGCTTCTAGTGTGGTGCACGTTAATAGCGTGGCTCATTTGTCGACCAACT 60  
Db |||||||  
Oy 106 ATCAAGGTGATGCTTCTAGTGTGGTGCACGTTAATAGCGTGGCTCATTTGTCGACCAACT 165  
Db |||||||

Oy 61 TCAACTTGTCCTAATAATAGCATCAGCTTTGATGCTGCAAAATGCCACCATTAACAATAT 120  
Db |||||||

Oy 166 TCAACTTGTCCTAATAATAGCATCAGCTTTGATGCTGCAAAATGCCACCATTAACAATAT 225  
Db |||||||

Oy 121 GCCACCTTTATGGAATCTCTTCGTAATCAAGCGAAGAGTCCAAACTAAATGCTATGCC 180  
Db |||||||

Oy 226 GCCACCTTTATGGAATCTCTTCGTAATCAAGCGAAGAGTCCAAACTAAATGCTATGCC 285  
Db |||||||

Oy 181 ATACCAATGCTACTGATGATCAATTCGACCCCTTAAGTACTTATTTGGTTAAGCTCCAAAGT 240  
Db |||||||

Oy 286 ATACCAATGCTACTGATGATCAATTCGACCCCTTAAGTACTTATTTGGTTAAGCTCCAAAGT 345  
Db |||||||

Oy 241 GCAACCTTAAACACCATTAACATGCTGAGAGCAATTAACATTAACGTTATGCGGTAT 300  
Db |||||||

Oy 346 GCAACCTTAAACACCATTAACATGCTGAGAGCAATTAACATTAACGTTATGCGGTAT 405  
Db |||||||

Oy 301 TCTGATCCCTCAATGCCAATAGTGTGCTTACCATATATTTATGATATACAGCAC 360  
Db |||||||

Oy 406 TCTGATCCCTCAATGCCAATAGTGTGCTTACCATATATTTATGATATACAGCAC 465  
Db |||||||

Oy 361 GAACGCACTGATGGAAGATACCTTTTGTCTCAAGTTCTAGTCTCTGTTGCAATGTC 420  
Db |||||||

Oy 466 GAACGCACTGATGGAAGATACCTTTTGTCTCAAGTTCTAGTCTCTGTTGCAATGTC 525  
Db |||||||

Oy 421 ATTAACCTAATAGCTTATATCGGACCATCGCAAGAGAGAGCAAGTAACACTCAAGAA 480  
Db |||||||

Oy 526 ATTAACCTAATAGCTTATATCGGACCATCGCAAGAGAGAGCAAGTAACACTCAAGAA 585  
Db |||||||

Oy 481 CAAGTCCAATGGGAATTCGAATACACTCAGAGTGCACATTCGMAAATCTCTGGAGTTGAT 540  
Db |||||||

Oy 586 CAAGTCCAATGGGAATTCGAATACACTCAGAGTGCACATTCGMAAATCTCTGGAGTTGAT 645  
Db |||||||

Oy 541 TCATTCTCTGTAAACCTGAGGCTTTTCTACTGTTAGGCTCCCAATGGTTTCAGAG 600  
Db |||||||

Oy 646 TCATTCTCTGTAAACCTGAGGCTTTTCTACTGTTAGGCTCCCAATGGTTTCAGAG 705  
Db |||||||

Oy 601 GCAGCGGATTCAGATACATAGAGAACCACTCAAGACTAATTTTATAGAGCAATTCAC 660  
Db |||||||

Oy 706 GCAGCGGATTCAGATACATAGAGAACCACTCAAGACTAATTTTATAGAGCAATTCAC 765  
Db |||||||

Oy 661 CCTGATCCCAAGTAAATTAATTTGGAGAGAGTGGGCAAAATCTCTGGAGCAATTCAC 720  
Db |||||||

Oy 766 CTTGATCCCAAGTAAATTAATTTGGAGAGAGTGGGCAAAATCTCTGGAGCAATTCAC 825  
Db |||||||

Oy 721 AATGCCAAGATGGGCTTTTACCCAAACCACTTCAAGCTAGTGGATGGCAAGGTACCAAG 780  
Db |||||||

Oy 826 AATGCCAAGATGGGCTTTTACCCAAACCACTTCAAGCTAGTGGATGGCAAGGTACCAAG 885  
Db |||||||

Oy 781 TGGATAGTTCTTACAGTGGATGAATCAATCGTATGCTGGCACTCCCTTAAGTACGTTAT 840  
Db |||||||

Oy 886 TGGATAGTTCTTACAGTGGATGAATCAATCGTATGCTGGCACTCCCTTAAGTACGTTAT 945  
Db |||||||

Oy 841 GGAACCTGTCAGACAACTTACCAAAATGCCATGTTCTCAAGTTATTAATTTCTACTTAT 900  
Db |||||||

Oy 946 GGAACCTGTCAGACAACTTACCAAAATGCCATGTTCTCAAGTTATTAATTTCTACTTAT 1005  
Db |||||||

Oy 901 TATAATATATGCTAATCTTGGTATGCTATTTTGAAGGATCTT 943  
Db |||||||

1006 TATAATATATGCTAATCTTGGTATGCTATTTTGAAGGATCTT 1048

RESULT 3  
AX427704  
LOCUS

792 bp DNA linear PAT 20-JUN-2002

AX427704

Sequence 3 from Patent WO0233107.

ACCESSION AX427704.1 GI:21537816

VERSION AX427704.1

KEYWORDS

SOURCE

ORGANISM

Phytolacca americana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Caryophyllales; Caryophyllaceae; Phytolacca

1

Neelam, A., Atkinson, H. J., McPherson, M. J. and Thomas, C. J. R.

Plant cell death system

Patent: WO 0233107-A 3 25-APR-2002;

CAMBRIDGE ADVANCED TECH (GB)

FEATURES

Location/Qualifiers

1..792

/organism="Phytolacca americana"

/db\_xref="taxon:3527"

/note="Binding site for primer PS1BF"

1..3

/note="Initiation codon added via PCR primer"

complement(436..462)

/note="Binding site for primer PS1SR"

463..492

/note="Binding site for primer PS2BF"

612

/note="Nucleotide change from published sequence"

681..686

/note="Sequence replacing removed XbaI site"

complement(765..792)

/note="Binding site for primer PS2SR"

790..792

/note="Stop codon added via PCR primer"

268 a 163 c 151 g 212 t

BASE COUNT

ORIGIN

Query Match 78.0%; Score 737; DB 6; Length 792;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 787; Conservative 0; Mismatches 1; Indels 0; Gaps 0;







SOURCE  
ORGANISM  
Phytolacca americana  
Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Caryophyllales; Caryophyllaceae; Phytolaccaceae; Phytolacca.  
1  
REFERENCE  
AUTHORS  
Neelam, A., Atkinson, H.J., McPherson, M.J. and Thomas, C.J.R.  
TITLE  
Plant cell death system  
JOURNAL  
Patent: WO 0233107-A 7 25-APR-2002;  
CAMBRIDGE ADVANCED TECH (GB)  
FEATURES  
Location/Qualifiers  
1..333  
/organism="Phytolacca americana"  
/db\_xref="taxon:3527"  
misc\_feature  
1..3  
/note="Initiation codon added via PCR primer"  
misc\_feature  
331..333  
/note="Stop codon added via PCR primer"  
BASE COUNT  
106 a 63 c 76 g 88 t  
ORIGIN  
34.8%; Score 329; DB 6; Length 333;  
Best Local Similarity 100.0%; Pred. No. 2e-168;  
Matches 329; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 532 GGAGTTGATTCATCCCTGTAATAACTGAGGCTTTTCTACTGCTAGCCATCCAAATG 591  
Db 4 GGAGTTGATTCATCCCTGTAATAACTGAGGCTTTTCTACTGCTAGCCATCCAAATG 63  
Qy 592 GTTTCAGAGGAGCGCGATTCAAGTACATAGACCAAGTCAAGACTAATTTTAATAGA 651  
Db 64 GTTTCAGAGGAGCGCGATTCAAGTACATAGACCAAGTCAAGACTAATTTTAATAGA 123  
Qy 652 CATCTCTACCTGATCCCAAGTAATTAATTGGAGGAGAGTGGGGCAAAATCTCTGAG 711  
Db 124 GCATTTACCTGATCCCAAGTAATTAATTGGAGGAGAGTGGGGCAAAATCTCTGAG 183  
Qy 712 GCAATTCACATGCCAAGTGGGCTTTTACCAAAACCACCTTGAGCTAGTGATGCCAAA 771  
Db 184 GCAATTCACATGCCAAGTGGGCTTTTACCAAAACCACCTTGAGCTAGTGATGCCAAA 243  
Qy 772 GGTACCAAGTGGATGTTCTTAGAGTGGATGAATCAATCTGATGTGGCACTCCTTAAG 831  
Db 244 GGTACCAAGTGGATGTTCTTAGAGTGGATGAATCAATCTGATGTGGCACTCCTTAAG 303  
Qy 832 TACGTTAATGAACCTGTGCAGACACTTA 860  
Db 304 TACGTTAATGAACCTGTGCAGACACTTA 332  
RESULT 8  
AX427716  
LOCUS  
AX427716  
DEFINITION  
Sequence 15 from Patent WO0233107.  
ACCESSION  
AX427716  
VERSION  
AX427716.1 GI:21537825  
KEYWORDS  
synthetic construct.  
SOURCE  
ORGANISM  
Phytolacca americana  
REFERENCE  
AUTHORS  
Neelam, A., Atkinson, H.J., McPherson, M.J. and Thomas, C.J.R.  
TITLE  
Plant cell death system  
JOURNAL  
Patent: WO 0233107-A 15 25-APR-2002;  
CAMBRIDGE ADVANCED TECH (GB)  
FEATURES  
Location/Qualifiers  
1..42  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="PSXDF primer"  
15..20  
/note="Nucleotides modified to remove XbaI site"  
BASE COUNT  
12 a 12 c 9 g 9 t

ORIGIN  
Query Match  
Best Local Similarity 4.4%; Score 42; DB 6; Length 42;  
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 736 CCTTACCCAAACCACTTGAGCTAGCTAGTCCCAAGGTACC 777  
Db 1 CCTTACCCAAACCACTTGAGCTAGCTAGTCCCAAGGTACC 42  
RESULT 9  
AX427717/c  
LOCUS  
AX427717  
DEFINITION  
Sequence 16 from Patent WO0233107.  
ACCESSION  
AX427717  
VERSION  
AX427717.1 GI:21537826  
KEYWORDS  
synthetic construct.  
SOURCE  
ORGANISM  
Phytolacca americana  
REFERENCE  
AUTHORS  
Neelam, A., Atkinson, H.J., McPherson, M.J. and Thomas, C.J.R.  
TITLE  
Plant cell death system  
JOURNAL  
Patent: WO 0233107-A 16 25-APR-2002;  
CAMBRIDGE ADVANCED TECH (GB)  
FEATURES  
Location/Qualifiers  
1..42  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="PSXDF primer"  
22..27  
/note="Nucleotides modified to remove XbaI site"  
BASE COUNT  
9 a 10 c 11 g 12 t  
ORIGIN  
Query Match  
Best Local Similarity 4.4%; Score 42; DB 6; Length 42;  
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 735 GCGTTTACCCAAACCACTTGAGCTAGCTAGTCCCAAGGTACC 776  
Db 42 GCGTTTACCCAAACCACTTGAGCTAGCTAGTCCCAAGGTACC 1  
RESULT 10  
AB071854  
LOCUS  
AB071854  
DEFINITION  
Phytolacca americana pap1 gene for PAP-S1, partial cds.  
ACCESSION  
AB071854  
VERSION  
AB071854.1 GI:19570837  
KEYWORDS  
Phytolacca americana DNA.  
SOURCE  
ORGANISM  
Phytolacca americana  
REFERENCE  
AUTHORS  
Honjo, E. and Watanabe, K.  
TITLE  
Cloning of genomic DNA encoding two types of pokeweed antiviral  
protein in seeds, PAP-S1 and PAP-S2, and functional comparison of  
their recombinant proteins with other PAP isoforms  
JOURNAL  
Unpublished  
REFERENCE  
AUTHORS  
Watanabe, K. and Honjo, E.  
TITLE  
Direct Submission  
JOURNAL  
Submitted (19-SEP-2001) Keio University, Saga University,  
Department of Applied Biological Sciences, Honjo-machi, Saga city,  
Saga 840-8502, Japan (E-mail: watanabe@saga-u.ac.jp,  
Tel: 81-952-28-8774, Fax: 81-952-28-8774)  
FEATURES  
Location/Qualifiers  
1..783  
/organism="Phytolacca americana"

```

09978274
/db_xref="taxon:3527"
1.783
/feature="paps1"
<1>783
/feature="paps1"
/feature="mature pokeweed antiviral protein-S1"
/codon_start=1
/protein_id="BAB86349.1"
/db_xref="GI:19570838"
/translation="INTIFDAGNATINKYATFMESLRNEAKDPSLKCYCTOMLPNTN
STIKLLVRLQASLKTITMLLRNLYVMGSDPYDKRCRIHENDIKAGTEYSUVEN
TLCPSNPVAPKIPYNGLYPTLEKAGVTSRQVGLGIISSDGIKSGGGSFTEK
IEAKELLVAIQWSEARFKYIENQVTFNRDPSNDRVLDLEENKGIKSTAIHNSK
NGALPKPLKLNADGKFWILVRVDEIKPKDVLNLYVNGTCQAT"
BASE COUNT 272 a 159 c 153 g 199 t
ORIGIN
Query Match 4.4% Score 42; DB 8; Length 783;
Best Local Similarity 100.0%; Pred. No. 2.6e-11;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Caps 0;
QY 493 GGAATTCAACTACTGACGACGACATTTGGAATAATCTCTCGA 534
DB 418 GGAATTCAACTACTGACGACGACATTTGGAATAATCTCTCGA 459
RESULT 11
LOCUS E05033 2472 bp DNA linear PAT 29-SEP-1997
DEFINITION DNA encoding antiviral protein.
ACCESSION E05033
VERSION E05033.1 GI:2173227
KEYWORDS JP 1993137580-A/1.
SOURCE Phytolacca americana.
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Caryophyllales; Caryophyllales; Phytolaccaceae; Phytolacca.
REFERENCE 1 (bases 1 to 2472)
AUTHORS Kataoka,J., Habuka,N., Masuda,O., Miyano,M. and Koizumi,A.
TITLE NEW PROTEIN AND GENE CAPABLE OF CODING THE SAME
JOURNAL Patent: JP 1993137580-A 1 01-JUN-1993;
JAPAN TOBACCO INC
COMMENT OS Phytolacca americana
PN JP 1993137580-A/1
PD 01-JUN-1993
PF 20-NOV-1991 JP 1991329672
PI KATAOKA JIRO, HABUKA NORIYUKI, MASUDA OSAMU, MIYANO MASASHI,
PC KOIWAI AKIRA
PC C12N15/29,C12N15/73,C12P21/02//C12N1/21.(C12P21/02,C12R1/19),
PC (C12N1/21),
PC (C12R1/19);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC Feature is identified by similarity;
FH Location/Qualifiers
FT CAAT_signal 549..552
FT CAAT_signal 627..630
FT TATA_signal 845..850
FT sig_peptide 1014..1085
FT mat_peptide 1086..11895
FT /product='antiviral protein'
FT CDS 1014..1898
FT polyA_site 2130..2135.
FEATURES
source
1..2472
/organism="Phytolacca americana"
/db_xref="taxon:3527"
BASE COUNT 864 a 375 c 463 g 770 t
us-09-978-274a-1.oli.rge
06/02/2003
ORIGIN
Query Match 3.4% Score 32; DB 6; Length 2472;
Best Local Similarity 100.0%; Pred. No. 7.3e-06;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Caps 0;
QY 577 GTAGCCATCAATGCTTTTCAGAGGCGCGCG 608
DB 1587 GTAGCCATCAATGCTTTTCAGAGGCGCGCG 1618
RESULT 12
LOCUS PTCAPAP 2472 bp DNA linear PLN 01-FEB-2000
DEFINITION P. americana DNA for alpha-PAP(pokeweed antiviral protein).
complete cds.
ACCESSION D10600 D90537
VERSION D10600.1 GI:218010
KEYWORDS alpha-PAP; antiviral protein; pokeweed antiviral protein (PAP);
ribosome-inactivating protein.
SOURCE Phytolacca americana DNA.
ORGANISM Phytolacca americana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Caryophyllales; Caryophyllales; Phytolaccaceae; Phytolacca.
REFERENCE 1 (bases 1 to 2472)
AUTHORS Kataoka,J., Habuka,N., Masuda,C., Miyano,M. and Koizumi,A.
TITLE Isolation and analysis of a genomic clone encoding a pokeweed
antiviral protein
JOURNAL Plant Mol. Biol. 20 (5), 879-886 (1992)
MEDLINE 93099240
REFERENCE 2 (bases 1 to 2472)
AUTHORS Kataoka,J.
TITLE Direct Submission
JOURNAL Submitted (23-OCT-1991) Jiro Kataoka, Japan Tobacco Inc., Life
Science Research Laboratory, 6-2 Umegakoka, Midori-ku, Yokohama,
Kanagawa 227, Japan (Tel:045-972-5901, Fax:045-972-6205)
COMMENT Submitted (23-Oct-1991) to DDBJ by:
Jiro Kataoka
Life Science Research Laboratory
Japan Tobacco Inc.
6-2 Umegakoka, Midori-ku
Yokohama 227
Japan
Phone: 045-972-5901
Fax: 045-972-6205.
FEATURES
source
1..2472
/organism="Phytolacca americana"
/db_xref="taxon:3527"
549..552
845..850
1014..1898
/codon_start=1
/product="alpha-PAP(pokeweed antiviral protein)"
/protein_id="BAA01451.1"
/db_xref="GI:218011"
/translation="MKMMVVVVVMMMLSLILKPPSTWINTTIFHVNALINKYATFM
KSTHNOAKDPTLCYGIPIMLPNTNLTPKYLLVTLQDSSLTITLMLKNNLYVMGVAD
TYNKGKRYHTFXDTSNTTERNDVMTTLCNPSSRVCKNNYDSSYPALLEKVCGRPHSQ
VQLGLOIINSIGICKIVGVDSFTKTEAEELLVAIQWSEARFKYIENQVTFNRDPSNDRV
YPNKVLNLEESWKGKSTAIHNAKNGALTSPLKKNAGSKNVLKLVNLIETFWGLLK
YVNGKTCATYOSAMFPHL"
1014..1085
1086..11895
sig_peptide
mat_peptide
/polyA_signal 2130..2135
/product="alpha-PAP(pokeweed antiviral protein)"
BASE COUNT 864 a 375 c 463 g 770 t

```

OY 577 GTAGCCATCCAAATGGTTTCAGAGCCGCGC 608  
 DB 1587 GTAGCCATCCAAATGGTTTCAGAGCCGCGC 1618  
 RESULT 13  
 AX427714 52 bp DNA linear PAT 20-JUN-2002  
 LOCUS AX427714  
 DEFINITION Sequence 13 from Patent WO0233107.  
 ACCESSION AX427714  
 VERSION AX427714.1 GI:21537823  
 KEYWORDS  
 SOURCE synthetic construct.  
 ORGANISM synthetic construct.  
 REFERENCE 1  
 AUTHORS Neelam A., Atkinson, H.J., Mcpherson, M.J. and Thomas, C.J.R.  
 TITLE Plant cell death system  
 JOURNAL Patent: WO 0233107-A 13 25-APR-2002;  
 CAMBRIDGE ADVANCED TECH (GB)  
 FEATURES  
 source  
 1..52  
 Location/Qualifiers  
 /organism="synthetic construct"  
 /db\_xref="taxon:32630"  
 /note="PS2BF primer"  
 misc\_feature 1..19  
 /note="Introduced restriction sites"  
 BASE COUNT 15 a 10 c 13 g 14 t  
 ORIGIN  
 Query Match 3.2% Score 30; DB 6; Length 52;  
 Best Local Similarity 100.0%; Pred. No. 0.00011;  
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 532 GGAGTTGATTCATTCCTGTAATAACTGAG 561  
 DB 23 GGAGTTGATTCATTCCTGTAATAACTGAG 52  
 RESULT 14  
 AY049785 1164 bp mRNA linear PLN 23-OCT-2001  
 LOCUS AY049785  
 DEFINITION Phytolacca acinosa anti-viral protein PAP (PAP) mRNA, complete cds.  
 ACCESSION AY049785  
 VERSION AY049785.1 GI:16356654  
 KEYWORDS  
 SOURCE Phytolacca acinosa.  
 ORGANISM Phytolacca acinosa.  
 REFERENCE 1  
 AUTHORS Peng, X., Yuan, J. and Qiang, B.  
 TITLE Direct Submission  
 JOURNAL Submitted (27-JUL-2001) Department of Biochemistry, Institute of Basic Medical Science, Chinese Academy of Medical Sciences, 5 Dong Dan San Tiao, Beijing 100005, P.R. China  
 FEATURES  
 source  
 1..1164  
 /organism="Phytolacca acinosa"  
 /db\_xref="taxon:107615"  
 gene 1..1164  
 CDS 2..943  
 /gene="PAP"  
 /codon\_start=1  
 /product="anti-viral protein PAP"  
 /protein\_id="AAL15442.1"  
 /db\_xref="GI:16356655"  
 /translation="MKSMLVVTISWLLIAPTSTHVNNTIIVNGSTTISKVATFLDN  
 LRNEADPSLXCYGIPMLPNTNPVKYVLVELOGSNKKTITLMLRNNLYVMGYSDPF  
 DTNRCYHIFDISGTERODVETTLCPNPNRSVKNINVDYRPTLESKAGVKSRSQV

OLGIQILDSNIGKISGVTSTFEKTEAEFLVAIOMVEAARFYICNUVKTNNKRAFN  
 PNPKYLNLEETWCKISTALHUANKGVLPKPLELVDSAGAKWILKVIDIKPIVALNLNY  
 VGGSCQTYNONAMFPOLIMSTYINMYANGLDLEFGP  
 BASE COUNT 403 a 202 c 219 g 340 t  
 ORIGIN  
 Query Match 3.2% Score 30; DB 8; Length 1164;  
 Best Local Similarity 100.0%; Pred. No. 9.4e-05;  
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 914 CTAATCTTGGTGATCTATTTCAGGATTCCT 943  
 DB 912 CTAATCTTGGTGATCTATTTCAGGATTCCT 941  
 RESULT 15  
 PAPAP 1164 bp mRNA linear PLN 18-NOV-1991  
 LOCUS P.americana PAP gene for anti-viral protein.  
 DEFINITION X55383  
 ACCESSION X55383  
 VERSION X55383.1 GI:20421  
 KEYWORDS antiviral protein; cell wall protein; PAP gene; ribosome;  
 inactivating protein.  
 SOURCE Phytolacca americana.  
 ORGANISM Phytolacca americana.  
 REFERENCE 1 (bases 1 to 1164)  
 AUTHORS Antoniw, J.F.  
 TITLE Direct Submission  
 JOURNAL Submitted (05-NOV-1990) Antoniw J.F., AFRC Inst of Arable Crops Research, Dept. of Plant Pathology, Rothamsted Experimental Station, Harpenden, Hert, AL5 2JQ, UK  
 REFERENCE 2 (bases 1 to 1164)  
 AUTHORS Lin, Q., Chen, Z.C., Antoniw, J.F. and White, R.F.  
 TITLE Isolation and characterization of a cDNA clone encoding the anti-viral protein from Phytolacca americana  
 JOURNAL Plant Mol. Biol. 17 (4), 609-614 (1991)  
 MEDLINE 92003676  
 PUBMED 1912488  
 FEATURES  
 Location/Qualifiers  
 1..1164  
 /organism="Phytolacca americana"  
 /db\_xref="taxon:3527"  
 /tissue\_type="leaf"  
 /dev\_stage="vegetative"  
 gene 2..943  
 /gene="PAP"  
 CDS 2..943  
 /gene="PAP"  
 /codon\_start=1  
 /product="anti-viral protein"  
 /protein\_id="CAA39054.1"  
 /db\_xref="GI:20422"  
 /translation="MKSMLVVTISWLLIAPTSTHVNNTIIVNGSTTISKVATFLDN  
 LRNEADPSLXCYGIPMLPNTNPVKYVLVELOGSNKKTITLMLRNNLYVMGYSDPF  
 ETNRCYHIFDISGTERODVETTLCPNPNRSVKNINVDYRPTLESKAGVKSRSQV  
 OLGIQILDSNIGKISGVTSTFEKTEAEFLVAIOMVEAARFYICNUVKTNNKRAFN  
 PNPKYLNLEETWCKISTALHUANKGVLPKPLELVDSAGAKWILKVIDIKPIVALNLNY  
 VGGSCQTYNONAMFPOLIMSTYINMYANGLDLEFGP

BASE COUNT 409 a 200 c 216 g 339 t  
 ORIGIN  
 Query Match 3.1% Score 29; DB 8; Length 1164;  
 Best Local Similarity 100.0%; Pred. No. 0.00033;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 915 TAATCTTGGTGATCTATTTCAGGATTCCT 943  
 DB 913 TAATCTTGGTGATCTATTTCAGGATTCCT 941

Mon Jun 2 10:54:54 2003

Search completed: June 2, 2003, 04:16:54  
Job time : 2637 secs

us-09-978-274a-1.oli.rge

Page 9

06/02/2003

29

09978274

SRNT

Printed: 11/17/2004



(pieterdejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources ([http://bacpac.med.buffalo.edu/ordering\\_bac.htm](http://bacpac.med.buffalo.edu/ordering_bac.htm)) or from Research Genetics ([inforesgen.com](http://inforesgen.com)). BAC end Web Server: <http://www.htsc.washington.edu>  
 Plate: 1109 row: F column: 6  
 Seq primer: SP6  
 Class: BAC ends  
 High quality sequence stop: 360.

## FEATURES

Location/Qualifiers  
 1..360

/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="Plate=1109 Col-6 Row=F"  
 /clone\_lib="RPC1-11 Human Male BAC Library"  
 /sex="male"  
 /note="Vector: pBACe3.6; Site\_1: EcoRI; Site\_2: EcoRI;  
 Male blood DNA was isolated from one randomly chosen donor  
 and partially digested with a combination of EcoRI and  
 EcoRI Methylase. Size selected DNA was cloned into the  
 pBACe3.6 vector at EcoRI sites"

BASE COUNT 100 a 58 c 54 g 148 t  
 ORIGIN

Query Match 2.4% Score 23; DB 17; Length 360;

Best Local Similarity 100.0%; Pred. No. 0.47;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 242 CAACCTAAACATTACACTA 264

DB 50 CAACCTAAACATTACACTA 28

## RESULT 2

AZ551992

LOCUS AZ551992 621 bp DNA linear GSS 20-NOV-2000  
 DEFINITION RPC1-23-210L20.TJ RPC1-23 Mus musculus genomic clone RPC1-23-210L20  
 , DNA sequence.

ACCESSION AZ551992

VERSION 1

KEYWORDS GSS.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 621)  
 AUTHORS Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret  
 ,B., Levins,M., McGann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.  
 and Fraser,C.M.

TITLE Mouse BAC End Sequences from Library RPC1-23  
 JOURNAL Unpublished (1999)  
 COMMENT Other\_GSSs: RPC1-23-210L20.TV  
 Contact: Shaying Zhao  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: szhao@tigr.org

Clones are derived from the mouse BAC library RPC1-23. For BAC

library availability, please contact Pieter de Jong  
 (pieterdejong.med.buffalo.edu). Clones may be purchased from  
 BACPAC Resources ([http://bacpac.med.buffalo.edu/ordering\\_bac.htm](http://bacpac.med.buffalo.edu/ordering_bac.htm))  
 or from Research Genetics (<http://inforesgen.com>). BAC end page:  
[http://www.tigr.org/tdb/bac\\_ends/mouse/bac\\_end\\_intro.html](http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html)  
 Plate: 210 row: L column: 20  
 Seq primer: SP6  
 Class: BAC ends.

Location/Qualifiers

1..621

/organism="Mus musculus"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="RPC1-23-210L20"

## FEATURES

source

/clone\_lib="RPC1-23"

/sex="Female"

/lab\_host="DH10B"

/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site\_1:  
 EcoRI; Site\_2: EcoRI; Female C57BL/6J mouse kidney and/or  
 brain genomic DNA was isolated and partially digested  
 with a combination of EcoRI and EcoRI Methylase. Size  
 selected DNA was cloned into the pBACe3.6 vector at the  
 EcoRI sites. The ligation products were transformed into  
 DH10B electrocompetent cells (BRL Life Technologies)."

BASE COUNT 209 a 98 c 102 g 222 t

ORIGIN

Query Match

Best Local Similarity 2.4% Score 23; DB 17; Length 621;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 884 TTATAATTCTACTTATTAAT 906

DB 247 TTATAATTCTACTTATTAAT 269

## RESULT 3

AZ243043/c

LOCUS

DEFINITION RPC1-23-36G6.TJB RPC1-23 Mus musculus genomic clone RPC1-21-36G6,  
 DNA sequence.

ACCESSION AZ243043

VERSION AZ243043.1 GI:8556234

KEYWORDS GSS.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 171)  
 AUTHORS Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret  
 ,B., Levins,M., McGann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.  
 and Fraser,C.M.

TITLE Mouse BAC End Sequences from Library RPC1-23  
 JOURNAL Unpublished (1999)  
 COMMENT Other\_GSSs: RPC1-21-36G6.TV  
 Contact: Shaying Zhao  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: szhao@tigr.org

Clones are derived from the mouse BAC library RPC1-23. For BAC

library availability, please contact Pieter de Jong  
 (pieterdejong.med.buffalo.edu). Clones may be purchased from  
 BACPAC Resources ([http://bacpac.med.buffalo.edu/ordering\\_bac.htm](http://bacpac.med.buffalo.edu/ordering_bac.htm))  
 or from Research Genetics (<http://inforesgen.com>). BAC end page:  
[http://www.tigr.org/tdb/bac\\_ends/mouse/bac\\_end\\_intro.html](http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html)  
 Plate: 36 row: G column: 6  
 Seq primer: SP6  
 Class: BAC ends.

Location/Qualifiers

1..171

/organism="Mus musculus"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="RPC1-23-36G6"

/clone\_lib="RPC1-23"

/sex="Female"

/lab\_host="DH10B"

/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site\_1:  
 EcoRI; Site\_2: EcoRI; Female C57BL/6J mouse kidney and/or  
 brain genomic DNA was isolated and partially digested  
 with a combination of EcoRI and EcoRI Methylase. Size  
 selected DNA was cloned into the pBACe3.6 vector at the  
 EcoRI sites. The ligation products were transformed into  
 DH10B electrocompetent cells (BRL Life Technologies)."

## FEATURES

source

Location/Qualifiers

1..171

/organism="Mus musculus"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="RPC1-23-36G6"

/clone\_lib="RPC1-23"

/sex="Female"

/lab\_host="DH10B"

/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site\_1:  
 EcoRI; Site\_2: EcoRI; Female C57BL/6J mouse kidney and/or  
 brain genomic DNA was isolated and partially digested  
 with a combination of EcoRI and EcoRI Methylase. Size  
 selected DNA was cloned into the pBACe3.6 vector at the  
 EcoRI sites. The ligation products were transformed into  
 DH10B electrocompetent cells (BRL Life Technologies)."

BASE COUNT 209 a 98 c 102 g 222 t

ORIGIN

Query Match

Best Local Similarity 2.4% Score 23; DB 17; Length 621;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 884 TTATAATTCTACTTATTAAT 906

DB 247 TTATAATTCTACTTATTAAT 269

BASE COUNT 35 a 67 c 23 g 46 t

Query Match 2.1% Score 20: DB 17: Length 171;  
Best Local Similarity 100.0% Pred. No. 20:  
Matches 20: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Oy 684 GGAGGAGAGTGGGGCAAA 703  
Db 153 GCAGGAGAGTGGGGCAAA 134

RESULT 4  
BE626842/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
JOURNAL  
COMMENT  
TITLE  
JOURNAL  
COMMENT  
FEATURES  
source

1. 333  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="IMAGE:476216"  
/sex="unknown"  
/tissue\_type="embryo"  
/dev\_stage="13.5-14.5dpc total fetus"  
/lab\_host="DH10B"  
/note="vector: pT7T3D-Pac (Pharmacia) with a modified  
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA  
was primed with a Not I - oligo(dT) primer [5].  
TGTATCAATGAGTGGAGCGCGCGGAAATTTTATTTTATTTT  
T 3'], on equal amounts of mRNA from 2 13.5dpc and 2  
14.5dpc embryos [total RNA provided by Minoru Ko, Wayne  
State Univ. from 2 ]; double-stranded cDNA was ligated to  
Eco RI adaptors (Pharmacia), digested with Not I and  
cloned into the Not I and Eco RI sites of the modified  
pT7T3 vector. Library went through one round of  
normalization, and was constructed by Bento Soares and  
M. Fatima Bonaldo."

BASE COUNT 114 a 58 c 62 g 99 t

Query Match 2.1% Score 20: DB 9: Length 333;  
Best Local Similarity 100.0% Pred. No. 21:  
Matches 20: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Oy 120 TGCACCTTTATGGAATCTC 139  
Db 320 TGCACCTTTATGGAATCTC 301

RESULT 5  
BE626842/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
JOURNAL  
COMMENT  
TITLE  
JOURNAL  
COMMENT  
FEATURES  
source

1. 356  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="IMAGE:3369199"  
/clone\_lib="Soares\_mammary\_gland\_MLMG"  
/sex="female (lactating)"  
/tissue\_type="mammary gland"  
/lab\_host="DH10B"  
/note="vector: pT7T3D-Pac (Pharmacia) with a modified  
polylinker; 1st strand cDNA was prepared from mammary  
gland tissue from a lactating female, and was then primed  
with a Not I - oligo(dT) primer. Double-stranded cDNA was  
ligated to Eco RI adaptors (Pharmacia), digested with Not  
I and cloned into the Not I and Eco RI sites of the  
modified pT7T3 vector. Library is normalized. Library  
was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 117 a 61 c 66 g 112 t

Query Match 2.1% Score 20: DB 10: Length 356;  
Best Local Similarity 100.0% Pred. No. 21:  
Matches 20: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Oy 120 TGCACCTTTATGGAATCTC 139  
Db 336 TGCACCTTTATGGAATCTC 317

RESULT 6  
U68743/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE

U68743  
U68743 Rittia pachyptilia anonymous nuclear DNA Rittia pachyptilia  
genomic clone pUKP113.1F, DNA sequence.  
U68743  
U68743 GI:1502556  
GSS.  
tube worm.  
Rittia pachyptilia  
Eukaryota; Metazoa; Pogonophora; Vestimentifera; Axonobranchia;  
Rittia; Rittidae; Rittia.  
1 (bases 1 to 359)  
Karl, S.A. and Wawrik, R.  
Identification of potentially functional regions in anonymous

BASE COUNT 35 a 67 c 23 g 46 t

Query Match 2.1% Score 20: DB 17: Length 171;  
Best Local Similarity 100.0% Pred. No. 20:  
Matches 20: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Oy 684 GGAGGAGAGTGGGGCAAA 703  
Db 153 GCAGGAGAGTGGGGCAAA 134

RESULT 4  
BE626842/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
JOURNAL  
COMMENT  
TITLE  
JOURNAL  
COMMENT  
FEATURES  
source

1. 333  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="IMAGE:476216"  
/sex="unknown"  
/tissue\_type="embryo"  
/dev\_stage="13.5-14.5dpc total fetus"  
/lab\_host="DH10B"  
/note="vector: pT7T3D-Pac (Pharmacia) with a modified  
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA  
was primed with a Not I - oligo(dT) primer [5].  
TGTATCAATGAGTGGAGCGCGCGGAAATTTTATTTTATTTT  
T 3'], on equal amounts of mRNA from 2 13.5dpc and 2  
14.5dpc embryos [total RNA provided by Minoru Ko, Wayne  
State Univ. from 2 ]; double-stranded cDNA was ligated to  
Eco RI adaptors (Pharmacia), digested with Not I and  
cloned into the Not I and Eco RI sites of the modified  
pT7T3 vector. Library went through one round of  
normalization, and was constructed by Bento Soares and  
M. Fatima Bonaldo."

BASE COUNT 114 a 58 c 62 g 99 t

Query Match 2.1% Score 20: DB 9: Length 333;  
Best Local Similarity 100.0% Pred. No. 21:  
Matches 20: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Oy 120 TGCACCTTTATGGAATCTC 139  
Db 320 TGCACCTTTATGGAATCTC 301

RESULT 5  
BE626842/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
JOURNAL  
COMMENT  
TITLE  
JOURNAL  
COMMENT  
FEATURES  
source

1. 356  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="IMAGE:3369199"  
/clone\_lib="Soares\_mammary\_gland\_MLMG"  
/sex="female (lactating)"  
/tissue\_type="mammary gland"  
/lab\_host="DH10B"  
/note="vector: pT7T3D-Pac (Pharmacia) with a modified  
polylinker; 1st strand cDNA was prepared from mammary  
gland tissue from a lactating female, and was then primed  
with a Not I - oligo(dT) primer. Double-stranded cDNA was  
ligated to Eco RI adaptors (Pharmacia), digested with Not  
I and cloned into the Not I and Eco RI sites of the  
modified pT7T3 vector. Library is normalized. Library  
was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 117 a 61 c 66 g 112 t

Query Match 2.1% Score 20: DB 10: Length 356;  
Best Local Similarity 100.0% Pred. No. 21:  
Matches 20: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Oy 120 TGCACCTTTATGGAATCTC 139  
Db 336 TGCACCTTTATGGAATCTC 317

RESULT 6  
U68743/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE

U68743  
U68743 Rittia pachyptilia anonymous nuclear DNA Rittia pachyptilia  
genomic clone pUKP113.1F, DNA sequence.  
U68743  
U68743 GI:1502556  
GSS.  
tube worm.  
Rittia pachyptilia  
Eukaryota; Metazoa; Pogonophora; Vestimentifera; Axonobranchia;  
Rittia; Rittidae; Rittia.  
1 (bases 1 to 359)  
Karl, S.A. and Wawrik, R.  
Identification of potentially functional regions in anonymous





db 325 TCCACCTTTATGGAATCTC 306

## RESULT 9

A1846009/c

## LOCUS

DEFINITION

UI-M-API-agl-e-12-0-UI-3' mRNA linear EST 15-JUL-1999

UI-M-API-agl-e-12-0-UI 3', mRNA sequence.

## ACCESSION

VERSION

## KEYWORDS

SOURCE

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;

1 (bases 1 to 406)

Bonaldo,M.F., Lennon,G. and Soares,M.B.

Normalization and subtraction: two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

9704477

Contact: Chin, H

National Institute of Mental Health

6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD

20892-9643, USA

Tel: 301 443 1706

Fax: 301 443 9890

Email: mestr@mail.nih.gov

The sequence contained an oligo-dT track that was present in the

oligonucleotide that was used to prime the synthesis of first

strand cDNA and therefore this may represent a bonafide poly A

tail. The sequence tag present in the cDNA between the NotI site

and the oligo-dT track served to verify it as a clone from the

normalized corpus striatum library cDNA Library Preparation: M.B.

Soares Lab Clone distribution: NIH BMAP cDNA clones will be made

available by the means that is soon to be determined. When NIH

determines the means for distribution of the BMAP cDNA clones, this

record will be updated accordingly when that means is determined.

Seq primer: M13 Forward

POLYA=Yes.

## FEATURES

source

Location/Qualifiers

1..406

/organism="Mus musculus"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="UI-M-API-agl-e-12-0-UI"

/clone.lib="NIH\_BMAP\_MST\_N"

/dev\_stage="27-32 days"

/lab\_host="DH10B (Life Technologies)"

/note="vector: pT73D-Pac (Pharmacia) with a modified

polylinker; Site\_1: Not I; Site\_2: Eco RI; The

NIH\_BMAP\_MST\_N library is a normalized library constructed

from mouse striatum. The tag is a string of 5 nucleotides

present between the Not I site and the oligo-dT track.

The library was constructed as described by Bonaldo,

Lennon and Soares, Genome Research 6: 791-806, 1996.

Tissue provided by Ms. Annie Novakovich, Zivic-Miller

Laboratories.

TAG\_LIB=NIH\_BMAP\_MST\_N

TAG\_TISSUE=corpus-striatum

TAG\_SEQ=ACGGC

BASE COUNT

138 a 75 c 84 g 109 t

## ORIGIN

Query Match

Best Local Similarity 100.0%; Pred. No. 21;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

## RESULT 10

BQ299989

## LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

## ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 410)

Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,

Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,

Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare

,M.J., Soares,F., Brentani,M.R., Reis,L.F., de Souza,S.J. and

Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

PROC. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MK12-MR1-KTU058-

191200-007-c10st3-2000-12-19kt4-1)

Seq primer: puc 18 forward

High quality sequence start: 3

High quality sequence stop: 41.

Location/Qualifiers

1..410

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone.lib="KT0058"

/dev\_stage="Adult"

/note="Organ: bladder\_tumor; Vector: puc18; Site\_1: SmaI;

Site\_2: SmaI; A mini-library was made by cloning products

derived from ORESTES PCR (U.S. Letters Patent application

No. 196,716 - Ludwig Institute for Cancer Research)

profiles into the puc 18 vector. Reverse transcription of

tissue mRNA and cDNA amplification were performed under

low stringency conditions."

BASE COUNT

100 a 111 c 101 g 98 t

## ORIGIN

Query Match

Best Local Similarity 100.0%; Pred. No. 21;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 81 GATCACCCTTTGATGCTGGAA 100

|||||

Db 297 GATCACCCTTTGATGCTGGAA 316

## RESULT 11

A2066860/c

## LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

|||||

325 TCCACCTTTATGGAATCTC 306

## RESULT 9

A1846009/c

## LOCUS

DEFINITION

UI-M-API-agl-e-12-0-UI-3' mRNA linear EST 15-JUL-1999

UI-M-API-agl-e-12-0-UI 3', mRNA sequence.

## ACCESSION

VERSION

## KEYWORDS

SOURCE

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;

1 (bases 1 to 406)

Bonaldo,M.F., Lennon,G. and Soares,M.B.

Normalization and subtraction: two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

9704477

Contact: Chin, H

National Institute of Mental Health

6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD

20892-9643, USA

Tel: 301 443 1706

Fax: 301 443 9890

Email: mestr@mail.nih.gov

The sequence contained an oligo-dT track that was present in the

oligonucleotide that was used to prime the synthesis of first

strand cDNA and therefore this may represent a bonafide poly A

tail. The sequence tag present in the cDNA between the NotI site

and the oligo-dT track served to verify it as a clone from the

normalized corpus striatum library cDNA Library Preparation: M.B.

Soares Lab Clone distribution: NIH BMAP cDNA clones will be made

available by the means that is soon to be determined. When NIH

determines the means for distribution of the BMAP cDNA clones, this

record will be updated accordingly when that means is determined.

Seq primer: M13 Forward

POLYA=Yes.

## FEATURES

source

Location/Qualifiers

1..406

/organism="Mus musculus"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="UI-M-API-agl-e-12-0-UI"

/clone.lib="NIH\_BMAP\_MST\_N"

/dev\_stage="27-32 days"

/lab\_host="DH10B (Life Technologies)"

/note="vector: pT73D-Pac (Pharmacia) with a modified

polylinker; Site\_1: Not I; Site\_2: Eco RI; The

NIH\_BMAP\_MST\_N library is a normalized library constructed

from mouse striatum. The tag is a string of 5 nucleotides

present between the Not I site and the oligo-dT track.

The library was constructed as described by Bonaldo,

Lennon and Soares, Genome Research 6: 791-806, 1996.

Tissue provided by Ms. Annie Novakovich, Zivic-Miller

Laboratories.

TAG\_LIB=NIH\_BMAP\_MST\_N

TAG\_TISSUE=corpus-striatum

TAG\_SEQ=ACGGC

BASE COUNT

138 a 75 c 84 g 109 t

## ORIGIN

Query Match

Best Local Similarity 100.0%; Pred. No. 21;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

|||||

325 TCCACCTTTATGGAATCTC 306

## RESULT 9

A1846009/c

## LOCUS

DEFINITION

UI-M-API-agl-e-12-0-UI-3' mRNA linear EST 15-JUL-1999

UI-M-API-agl-e-12-0-UI 3', mRNA sequence.

## ACCESSION

VERSION

## KEYWORDS

SOURCE

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;

1 (bases 1 to 406)

Bonaldo,M.F., Lennon,G. and Soares,M.B.

Normalization and subtraction: two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

9704477

Contact: Chin, H

National Institute of Mental Health

6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD

20892-9643, USA

Tel: 301 443 1706

Fax: 301 443 9890

Email: mestr@mail.nih.gov

The sequence contained an oligo-dT track that was present in the

oligonucleotide that was used to prime the synthesis of first

strand cDNA and therefore this may represent a bonafide poly A

tail. The sequence tag present in the cDNA between the NotI site

and the oligo-dT track served to verify it as a clone from the

normalized corpus striatum library cDNA Library Preparation: M.B.

Soares Lab Clone distribution: NIH BMAP cDNA clones will be made

available by the means that is soon to be determined. When NIH

determines the means for distribution of the BMAP cDNA clones, this

record will be updated accordingly when that means is determined.

Seq primer: M13 Forward

POLYA=Yes.

## FEATURES

source

Location/Qualifiers

1..406

/organism="Mus musculus"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="UI-M-API-agl-e-12-0-UI"

/clone.lib="NIH\_BMAP\_MST\_N"

/dev\_stage="27-32 days"

/lab\_host="DH10B (Life Technologies)"

/note="vector: pT73D-Pac (Pharmacia) with a modified

polylinker; Site\_1: Not I; Site\_2: Eco RI; The

NIH\_BMAP\_MST\_N library is a normalized library constructed

from mouse striatum. The tag is a string of 5 nucleotides

present between the Not I site and the oligo-dT track.

The library was constructed as described by Bonaldo,

Lennon and Soares, Genome Research 6: 791-806, 1996.

Tissue provided by Ms. Annie Novakovich, Zivic-Miller

Laboratories.

TAG\_LIB=NIH\_BMAP\_MST\_N

TAG\_TISSUE=corpus-striatum

TAG\_SEQ=ACGGC

BASE COUNT

138 a 75 c 84 g 109 t

## ORIGIN

Query Match

Best Local Similarity 100.0%; Pred. No. 21;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

|||||

325 TCCACCTTTATGGAATCTC 306

## RESULT 9

A1846009/c

## LOCUS

DEFINITION

UI-M-API-agl-e-12-0-UI-3' mRNA linear EST 15-JUL-1999

UI-M-API-agl-e-12-0-UI 3', mRNA sequence.

## ACCESSION

VERSION

## KEYWORDS

SOURCE

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;

1 (bases 1 to 406)

Bonaldo,M.F., Lennon,G. and Soares,M.B.

Normalization and subtraction: two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

9704477

Contact: Chin, H

National Institute of Mental Health

6001 Executive

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 428)  
Zhou, S., Nierman, W., Feldblyum, T., Malek, J., Shatsman, S., Akinret, B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Kroll, H., de Jong, P., and Fraser, C.M.  
Mouse BAC End Sequences from Library RPCI-23  
Unpublished (1999)  
Other GSSs: RPCI-23-433M3.TV  
Contact: Shaving Zhao  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: szhaod@tigr.org

Clones are derived from the mouse BAC library RPCI-23. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/orderingframe.htm>) or from Resea ch Genetics ([inforesgen.com](http://inforesgen.com)). BAC end page: [http://www.tigr.org/tldb/bac\\_ends/mouse/bac\\_end\\_intro.html](http://www.tigr.org/tldb/bac_ends/mouse/bac_end_intro.html)

plate: 433 row: M column: 3  
Seq primer: SP6  
Class: BAC ends.  
Location/Qualifiers  
1. 428  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="RPCI-23-433M3"  
/clone\_lib="RPCI-23"  
/sex="Female"  
/lab\_host="DH10B"  
/note="Organ: Kidney/Brain; Vector: pBAC3.6; Site 1: EcoRI; Site 2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBAC3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."  
82 a 141 c 76 g 129 t

# FEATURES

source  
1. 428  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="RPCI-23-433M3"  
/clone\_lib="RPCI-23"  
/sex="Female"  
/lab\_host="DH10B"  
/note="Organ: Kidney/Brain; Vector: pBAC3.6; Site 1: EcoRI; Site 2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBAC3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."  
82 a 141 c 76 g 129 t

# BASE COUNT

82 a 141 c 76 g 129 t

Query Match 2.1% Score 20; DB 17; Length 428;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 684 GGAGGAGAGTGGGCAAAA 703  
Db 172 GGAGGAGAGTGGGCAAAA 153  
|||||

# RESULT 12

BE653559  
LOCUS 474 bp mRNA linear EST 06-SEP-2000  
DEFINITION UI-M-AP1-agi-e-12-0-UI\_r1 NIH\_BMAP\_MST\_N Mus musculus cDNA clone  
UI-M-AP1-agi-e-12-0-UI 5', mRNA sequence.  
ACCESSION BE653559  
VERSION BE653559.1 GI:9979466  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 474)  
Bonaldo, M.F., Lennon, G. and Soares, M.B.  
Normalization and subtraction: two approaches to facilitate gene discovery  
Genome Res. 6 (9), 791-806 (1996)  
9704477  
Contact: Chin, H  
National Institute of Mental Health

6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD 20892-9643, USA  
Tel: 301 443 1706  
Fax: 301 443 9890  
Email: mst@mail.nih.gov  
cDNA Library Preparation: M.B. Soares Lab Clone Distribution:  
Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It should be noted that Bento Soares is generating a small number of additional specialized non-redundant arrays of BMAP cDNAs whose availability will be considered under appropriate and limited collaborative arrangements  
Seq primer: M13 Reverse.  
Location/Qualifiers  
1. 474  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UI-M-AP1-agi-e-12-0-UI"  
/clone\_lib="NIH\_BMAP\_MST\_N"  
/dev\_stage="27-32 days"  
/lab\_host="DH10B (Life Technologies)"  
/note="Vector: p773D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; The NIH\_BMAP\_MST\_N library is a normalized library constructed from mouse striatum. The tag is a string of 5 nucleotides present between the Not I site and the oligo-dT track. The library was constructed as described by Bonaldo, Lennon and Soares, Genome Research 6: 791-806, 1996. Tissue provided by Ms. Annie Novakovich, Zivir-Miller Laboratories."  
104 a 160 c 95 g 115 t

# FEATURES

source  
1. 474  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UI-M-AP1-agi-e-12-0-UI"  
/clone\_lib="NIH\_BMAP\_MST\_N"  
/dev\_stage="27-32 days"  
/lab\_host="DH10B (Life Technologies)"  
/note="Vector: p773D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; The NIH\_BMAP\_MST\_N library is a normalized library constructed from mouse striatum. The tag is a string of 5 nucleotides present between the Not I site and the oligo-dT track. The library was constructed as described by Bonaldo, Lennon and Soares, Genome Research 6: 791-806, 1996. Tissue provided by Ms. Annie Novakovich, Zivir-Miller Laboratories."  
104 a 160 c 95 g 115 t

# Query Match

2.1% Score 20; DB 10; Length 474;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

# Qy

120 TGGCACCCTTTATGGAATCTC 139  
|||||

# Db

380 TGGCACCCTTTATGGAATCTC 399  
|||||

# RESULT 13

BM124329/c  
LOCUS 486 bp mRNA linear EST 01-FEB-2002  
DEFINITION L0538E12-3 NIA Mouse Newborn Heart cDNA Library Mus musculus cDNA clone L0538E12 3', mRNA sequence.  
ACCESSION BM124329  
VERSION BM124329.1 GI:17108097  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 486)  
Piao, Y., Kargul, G.J., Dudekula, D.B., Olan, Y., Pantano, S., Lim, M.K. and Ko, M.S.H.  
Systematic Analyses of NIA Mouse Newborn Heart cDNA Library  
Unpublished (2001)  
Contact: Dawood B. Dudekula  
Laboratory of Genetics  
National Institute on Aging/National Institutes of Health  
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
Email: edna@lysun-grc.nia.nih.gov  
Plate: L0538 row: E column: 12  
Seq primer: -21M13 Forward  
High quality sequence stop: 486  
polyA-Yes.  
Location/Qualifiers  
1. 486  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="niaEST:L0538E12-3"

```

/db_xref="taxon:10090"
/clone_lib="NIA Mouse Newborn Heart cDNA Library"
/tissue_type="Newborn Heart"
/dev_stage="Newborn"
/lab_host="DH108"
/notes="Organ: heart; vector: pSPORT1 (Invitrogen); Site_1:
SalI; Site_2: NotI; Mouse cDNA project by the Laboratory
of Genetics, National Institute on Aging (NIA), Intramural
Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA).
Double-stranded cDNAs were synthesized with an Oligo(dT)
primer (Invitrogen, 5'-
GACTAGTCTAGATCGGAGCGCGCCCTTTT-3') from
24.9 microgram of total RNA, treated with RNase
polymerase, and purified by ethanol-precipitation. The
cDNAs were ligated to Loe-linker LL-Sal3 (Ref.
Development 127:1737-1749 (2000) PMID: 107252491),
purified by phenol/chloroform, and separated from free
linkers by Centricon 100. Then, the cDNAs were digested
with SalI and NotI enzymes, and cloned into SalI and NotI
site of pSPORT1 plasmid vector. The DH108 E. coli host was
transformed with ligation mixture by the chemical method.
The average insert size is about 1.8 kb. The library was
constructed by Yulan Piao (NIA)."
BASE COUNT      153 a   97 c   114 g   122 t
ORIGIN
Query Match      2.1%; Score 20; DB 13; Length 486;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 120 TGCACCTTTATGGAATCTC 139
      |||||||
Db 325 TGCACCTTTATGGAATCTC 306

RESULT 14
AA058263 489 bp mRNA linear EST 19-SEP-1996
clone IMAGE:480339 5', mRNA sequence.
AA058263
AA058263.1 GI:1551098
EST.
house mouse.
ORGANISM
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 489)
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.
The WashU-HMNI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMNI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL: contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MG1:291083
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 464.
Location/Qualifiers
1..489
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"

```

```

/clone="IMAGE:480339"
/clone_lib="Soares mouse embryo NBMEL3.5 14.5"
/sex="unknown"
/tissue_type="embryo"
/dev_stage="13.5-14.5dpc total fetus"
/lab_host="DH108"
/notes="vector: pTT730-pac (Pharmacia) with a modified
polylinker; Site_1: NotI; Site_2: EcoRI; 3st strand cDNA
was primed with a NotI - Oligo(dT) primer 15'-
TGTACATCTCAACTGGAGCGCGCGGAAATTTT-3'
T 3', on equal amounts of mRNA from 2 14.5dpc and 2
14.5dpc embryos [total RNA provided by Minou K. Wayne
State Univ., from 2 1; double-stranded cDNA was ligated to
EcoRI adaptors (Pharmacia), digested with NotI and
cloned into the NotI and EcoRI sites of the modified
pTT730 vector. Library went through one round of
normalization, and was constructed by Rento Soares and
M. Fatima Bonaldo."
BASE COUNT      92 a   176 c   102 g   119 t
ORIGIN
Query Match      2.1%; Score 20; DB 9; Length 489;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 120 TGCACCTTTATGGAATCTC 139
      |||||||
Db 466 TGCACCTTTATGGAATCTC 485

RESULT 15
AA0512831 498 bp DNA linear GSS 05-OCT-2000
clone UUCIM0358K09 R, DNA sequence.
AA0512831
AA0512831.1 GI:10694147
GSS.
house mouse.
ORGANISM
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 498)
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Humil, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Kelly,
M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.
and Wright, D., Weiss, R.
Mouse whole genome scaffolding with paired end reads from 10Kb
plasmid inserts
Unpublished (2000)
Contact: Robert H. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: dduval@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0358 row: K column: 09
Seq primer: CACACAGGAACACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 498.
Location/Qualifiers
1..498
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_lib="Mouse 10Kb plasmid UUCIM library"
/clone="UUCIM0358K09"
/sex="Male"
/lab_host="E. coli strain XL10-Gold, TI-resistant, F-"
/notes="Vector: PWD42nv; Purified genomic DNA from M.

```

*musculus* C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (g14732114|g1AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance.

96 c	91 g	161 t
------	------	-------

BASE COUNT	150 a	96 c	91 g	161 t
ORIGIN				

BASE CO  
ORIGIN

```

Query Match      2.18; Score 20; DB 17; Length 498;
Best Local Similarity 100.00; Pred. No. 22;

```

Query Match 2.1%; Score 20; DB  
Best Local Similarity 100.0%; Pred. No. 22;

```

Best local similarity 100.0%, P1ED. NO. 22,
Matches 20; Conservative 0; Mismatches 0; Indels 0; Caps 0;

```

Qy 249 AAAACCATTA CACTAATGC 268

Dbb 397 AAAAACCATTACACTAATGC 416

Search completed: June 2, 2003, 04:50:14  
Job time : 1679 secs

```

search completed: Jun
Job time : 1679 secs

```

Printed: 11/17/2004

SRNT

06/02/2003

please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Uosegawa and Aaron Mammeter in Peter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

```
FEATURES
  source
    1..1001
    /organism="Drosophila melanogaster"
    /db_xref="taxon:7227"
    /clone="BACR15H24"
    /clone_lib="RPCI-98"
    /note="end : TET3"
  BASE COUNT   159 a   148 c   151 g   212 t   331 others
  ORIGIN
    166 CTAATGCTATCGCATACCAATGCTACCTGATGATTAATTCGACCCCTAAGTACTATTG 225
    Db  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
    86 MMTAAAMWTTTMAAAAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAW 145
    Qy  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
    226 GTTAAGCTCCAGCTGCAACCTTAAACCACTTACACTATGCTGAGACCAATACTTA 285
    Db  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
    146 SMWTTMAAAAMSRWTTWRCAMCAAMCHWTWACHATWATWAAWMAWMAWMAWMAW 205
    Qy  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
    286 TAGCTGATGGCTATCTGATGCTCCTCAATGGAATGCTGCTGCTGCTGCTGCTGCT 345
    Db  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
    206 MYTTTGGGGYAAWTTMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMW 265
    Qy  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
    346 GATATTACAAGCAGCAAGCAGCTGATGAGCAATCTCTTGTCTCAAGTCTAGTTCT 405
    Db  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
    266 AAWYCSGGYCCMAGCHMKWGGYWKWGGGGGAATWCTTTCGGTTCWYCCMACCCC 325
    Qy  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
    406 CGTGTTCATGTCATTAACATACTATATCCGACCATGGAAGAAAGACGAGAA 465
    Db  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
    326 MAWYCCCKKAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMA 385
    Qy  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
    466 GTAAACTCAGAATCAAGTCCCAATGCGAATCAATCAATCAATCAATCAATCAAT 525
    Db  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
    386 KYCVTCCCAAYCCAMMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAW 445
    Qy  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
    526 ATCTCTGGAGTTCATTCATTCCTGTAAACACTGAGGCTTTTCTACTGTGAGCATC 585
    Db  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
    446 TYCHMMWKGTTTCATTTAAWGGGCGCMWTTTTRRRGGCTTTTTCATTCCTTTC 505
    Qy  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
    586 CAAATGGTTTCAGAGGAGCGGATCAAGTACATAGAGAACCAAGTCAAGACTATTT 645
    Db  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
    506 WTTTAAWTTTGTAWTTTGTGSCWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAW 565
    Qy  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
    646 AATAGACCATTC 657
    Db  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
    566 TTTAAATTTTTC 577
```

```
RESULT 15
AW230870
LOCUS
DEFINITION
  uo68h01.y1 NCI_CGAP_Mam1 Mus musculus cdna clone IMAGE:2647729 5'
  similar to SW:KG3B_HUMAN P49841 GLYCOGEN SYNTHASE KINASE-3 BETA ;
  mRNA sequence.
ACCESSION
  AW230870
VERSION
  AW230870.1 GI:6560166
KEYWORDS
  EST.
SOURCE
  house mouse.
  ORGANISM
    Mus musculus
```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 557)  
NCI\_CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
Tumor Gene Index  
Unpublished (1997)  
Other ESTs: uo68h01.x1  
Contact: Robert Strausberg, Ph.D.  
Email: [cgaps-r@mail.nih.gov](mailto:cgaps-r@mail.nih.gov)  
Tissue Procurement: Gilbert Smith, Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The J.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone Distribution: NCI-CGAP clone distribution information can be found through the J.M.A.G.E. Consortium/LLNL at:  
[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)

MGI:1028181  
Seq primer: -40RP from Gibco  
High quality sequence stop: 347.  
Location/Qualifiers  
1..557  
/organism="Mus musculus"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:2647729"  
/clone\_lib="NCI\_CGAP\_Mam1"  
/tissue\_type="tumor, biopsy sample"  
/dev\_stage="10 months, virgin"  
/lab\_host="DH10B"  
/note="Organ: mammary; Vector: pCMV-SPORT6; Site: 1; Salt: Site: 2; NotI: Cloned unidirectionally. Primer: Glluo dt. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"  
BASE COUNT 160 a 120 c 147 g 130 t  
ORIGIN

Query Match 4.3%; Score 40.4; DB 10; Length 557;  
Best Local Similarity 54.8%; Pred. No. 8.5;  
Matches 80; Conservative 0; Mismatches 66; Indels 0; Gaps 0;  
Qy 496 ATTCAATATCTCAGCAGTGACATTCGAAAATCTCTGGAGTTGATTTCCTCTTAAA 555  
Db :  
192 AGTTATACAGACACGAAAGTGAATCGAATGATCATTTGGTGTGATATCAAGCAAA 251  
Qy 556 ACTGAGGCTTTTCTTCTACTGTAGCCATCCAAATGTTTTCAGAGCCAGCCGATTC 615  
Db :  
252 CTTTGTCTTCTGAGAACTGTTGCCATCAAGAAAGTTCTACAGCACAAGCGATTTAAG 311  
Qy 616 TACATAGACAACCAAGTCAAGACTAA 641  
Db :  
312 ACCGAGAGCTCCAGATCATAGAAA 337

Search completed: June 2, 2003, 02:32:40  
Job time : 1678 secs







FORWARD: T7  
BACKWARD: T3  
Plate: L48-5 row: G column: 12  
Seq primer: T3  
High quality sequence stop: 350  
POLYA=No.

FEATURES  
Source  
Location/Qualifiers  
1. .489  
/organism="Mesembryanthemum crystallinum"  
/db\_xref="taxon:3544"  
/clone="L48-484"  
/clone\_lib="Ice plant Lambda Uni-Zap XR expression library  
48 hours NaCl treatment"  
/tissue\_type="Leaf, 48 h 0.4M NaCl"  
/dev\_stage="Six week old"  
/note="Vector: Lambda Uni-Zap XR, Bluescript SK-; Site\_1:  
ECORI; Site\_2: XhoI"  
BASE COUNT 160 a 98 c 106 g 125 t  
ORIGIN

Query Match 4.9% Score 46.6; DB 10; Length 489;  
Best Local Similarity 54.7%; Pred. No. 0.2;  
Matches 122; Conservative 0; Mismatches 89; Indels 12; Gaps 1;

QY 113 ACAATATGCCACCTTTATGGAACTCTTCGTAATCAAGCGAAGATCCAAACTAAAT 172  
Db 113 AATATATGCCACCTTTATGGAACTCTTCGTAATCAAGCGAAGATCCAAACTAAAT 172  
Db 196 ATACATACCATACATCTCCCTGGAATCTACGCAATCAACTGAAGGATCCCAATGACAAC 255  
QY 173 GCTATGGCATCAATGCTACCTGATGATCAATTCGACCCC-----TAAGTACT 220  
Db 173 GCTATGGCATCAATGCTACCTGATGATCAATTCGACCCC-----TAAGTACT 220  
Db 256 CCATAGTATATACGATGCGCAAAACCTACGACTTCCAAAGATGGACAACGATACATATC 315  
QY 221 TATTGGTTAAGCTCCAAAGTGGCAAACTAAAAACCATTAACATATGCTGAGACGAAATA 280  
Db 221 TATTGGTTAAGCTCCAAAGTGGCAAACTAAAAACCATTAACATATGCTGAGACGAAATA 280  
Db 316 TGTACTAGAGCTGCGAGCGAGCAGCACCACCAAGGTCATATATCTCTAGACAGAAGCA 375  
QY 281 ACTTATACGTGATGGCTATCTTCGATCCCTTCAATGGCAATA 323  
Db 281 ACTTATACGTGATGGCTATCTTCGATCCCTTCAATGGCAATA 323  
Db 376 ATTTATTTTCATGGCTACAGTGACATGGTATTGAAATAA 418

RESULT 9  
AA856221 267 bp mRNA linear EST 20-FEB-2001  
LOCUS L30-24273 ice plant Lambda Uni-Zap XR expression library, 30 hours  
DEFINITION NaCl treatment Mesembryanthemum crystallinum cDNA clone L30-242 5'  
similar to Ribosome inactivating protein precursor (antiviral  
protein), mRNA sequence.  
ACCESSION AA856221  
VERSION AA856221.1 GI:2944523  
KEYWORDS EST.  
SOURCE Mesembryanthemum crystallinum  
ORGANISM Common Iceplant.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Caryophyllales; Caryophyllales; Alzooceae; Mesembryanthemum.  
1 (bases 1 to 267)  
Cushman, J.C.  
An expressed sequence tag database for the common ice plant.  
Mesembryanthemum crystallinum  
Unpublished (1997).  
Contact: Cushman JC  
Department of Biochemistry  
University of Nevada  
MS200, Reno, NV 89557-0014, USA  
Tel: 775-784-1918  
Fax: 775-784-1650  
Email: jcushman@unr.edu  
PCR Primers  
FORWARD: T7  
BACKWARD: T3  
Plate: L30-3 row: D column: 10  
Seq primer: T3

us-09-978-274a-1.lst

High quality sequence stop: 230.  
Location/Qualifiers  
1. .267  
/organism="Mesembryanthemum crystallinum"  
/db\_xref="taxon:3544"  
/clone="L30-242"  
/clone\_lib="Ice plant Lambda Uni-Zap XR expression library  
30 hours NaCl treatment"  
/tissue\_type="Leaf, 30 h 0.4M NaCl"  
/dev\_stage="Six week old"  
/note="Vector: Lambda Uni-Zap XR, Bluescript SK-; Site\_1:  
ECORI; Site\_2: XhoI"  
BASE COUNT 85 a 48 c 69 g 65 t  
ORIGIN

Query Match 4.8% Score 45.4; DB 9; Length 267;  
Best Local Similarity 53.0%; Pred. No. 0.35;  
Matches 97; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 453 AAGAAGCAGAGTAAGTAACTCAAGAAATCAAGTCCCAATGGGAATTCAAATATCTACGAC 512  
Db 453 AAGAAGCAGAGTAAGTAACTCAAGAAATCAAGTCCCAATGGGAATTCAAATATCTACGAC 512  
Db 25 ACAGATCTCTCAACACGACGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 84  
QY 513 TGACATTTGGAATATCTCTGGAGTTGATTCATTCCTGTAATAAAGTGGGCTTTTCTTCT 572  
Db 513 TGACATTTGGAATATCTCTGGAGTTGATTCATTCCTGTAATAAAGTGGGCTTTTCTTCT 572  
Db 85 TCGCATCGAGTCGGTTTATGGTAAGACATCGACAGCAAGTCAAGCAAGTCAAGCAAGTCAAG 144  
QY 573 ACTGTTAGCATCCCAATGTTTTCAGACGCGCGGATTCAGTACATAGACAAATCAACT 632  
Db 573 ACTGTTAGCATCCCAATGTTTTCAGACGCGCGGATTCAGTACATAGACAAATCAACT 632  
Db 145 GCTGATGTCATCAGATGTTTCTGAGCAGCAGGATTCAGATATATTGAGAGTAAGCT 204  
QY 633 CAA 635  
Db 205 GAA 207

RESULT 10  
BE036639 1021 bp mRNA linear EST 07-JUN-2000  
LOCUS MP03H03 MP Mesembryanthemum crystallinum cDNA 5' similar to  
DEFINITION ribosome-inactivating protein gelonin, mRNA sequence.  
ACCESSION BE036639  
VERSION BE036639.1 GI:8331648  
KEYWORDS EST.  
SOURCE Mesembryanthemum crystallinum  
ORGANISM Common Iceplant.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Caryophyllales; Caryophyllales; Alzooceae; Mesembryanthemum.  
1 (bases 1 to 1021)  
Bohnert, H.J., Borchert, C., Brazillie, S., Brooks, J., Eaton, M., Ferreira, H., Kawasaki, S., McCullough, A., Michalowski, C.B., Palacios, C.,  
Scara, G., Wheeler, M. and Zepeda, G.R.  
Functional Genomics of Plant Stress Tolerance  
Unpublished (2000)  
Contact: Michalowski, C.B.  
University of Arizona  
Bio Sciences West room 513, Tucson, AZ 85721, USA  
Tel: 520-621-7982  
Fax: 520-621-1697  
Email: cbm@u.arizona.edu  
An open reading frame exists.  
Insert Length: 1 Std Error: 0.00.  
Location/Qualifiers  
1. .1021  
/organism="Mesembryanthemum crystallinum"  
/db\_xref="taxon:3544"  
/clone\_lib="Mp"  
/tissue\_type="apical meristem and leaf primordia"  
/dev\_stage="6 weeks"  
/note="3 d 500mM NaCl"  
BASE COUNT 302 a 242 c 235 g 241 t 1 others  
ORIGIN

KEYWORDS EST.  
SOURCE common iceplant.  
ORGANISM Mesembryanthemum crystallinum

REFERENCE Bohnert.H.J., Borchert,C., Brazille,S., Brooks,J., Eaton,M., Ferrea  
AUTHORS H., Kawasaki,S., McColough,A., Michalowski,C.B., Palacio,C.,  
TITLE Functional Genomics of Plant Stress Tolerance  
JOURNAL Unpublished (2000)  
COMMENT Contact: Michalowski,C.B.  
University of Arizona  
Bio Sciences West room 513, Tucson, AZ 85721, USA  
Tel: 520-621-7982  
Fax: 520-621-1697  
Email: cbm@u.arizona.edu.

FEATURES source  
1. 639  
Location/Qualifiers  
/organism="Mesembryanthemum crystallinum"  
/db\_xref="taxon:3544"  
/clone\_lib="Mp"  
/tissue\_type="apical meristem and leaf primordia"  
/dev\_stage="6 weeks"  
/note="3 d 500mM NaCl"  
BASE COUNT 226 a 108 c 146 g 159 t  
ORIGIN

Query Match 5.48; Score 50.8; DB 10; Length 639;  
Best Local Similarity 61.2; Pred. No. 0.018;  
Matches 104; Conservative 0; Mismatches 57; Indels 9; Gaps 1;  
QY 463 GAAGTAACCAAGAAATCAAGTCCCAATGGGAAATCAATCACTACGACGATGCA 522  
DB 308 GGAGCTAAGGAGAAAGAGGAGGAGGATGATATTTAGTAACTAATATTCATATA 367  
QY 523 AAAATCTCTGGAGTTGATTCATCCCTGTAAGAACTGAGGCTTTTTCATGAGTAC 502  
DB 368 AACAGC-----GATAAAATCAATGAAAGAGTACAGGCTACATTTCTTCTAGTTC 418  
QY 583 ATCCAAATGTTTCAGAGGAGGAGGAGGATTCAGATACATAGACACCAACT 632  
DB 419 ATTCAAATGTTTCAGAGGAGGAGGAGGATTCATACATAGAGGAGGAGT 468

RESULT 7  
AW053634  
LOCUS 837 bp mRNA linear EST 20-FEB-2001  
DEFINITION L30-140173 Ice plant Lambda Uni-Zap XR expression library, 30 hours  
NaCl treatment Mesembryanthemum crystallinum cDNA clone L30-1401  
5', mRNA sequence.  
ACCESSION AW053634  
VERSION AW053634.1 GI:5916827  
KEYWORDS EST.  
SOURCE common iceplant.  
ORGANISM Mesembryanthemum crystallinum

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
AUTHORS Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
TITLE Caryophyllales; Alzooceae; Mesembryanthemum.  
JOURNAL 1 (bases 1 to 837)  
COMMENT Cushman, J.C.  
An expressed sequence tag database for the common ice plant,  
Mesembryanthemum crystallinum  
Unpublished (1997)  
Contact: Cushman JC  
Department of Biochemistry  
University of Nevada  
MS200, Reno, NV 89557-0014, USA  
Tel: 775-784-1918  
Fax: 775-784-1650  
Email: jcushman@unr.edu  
Poly(A) tail, 10 nt: 838..847

PCR Primers  
FORWARD: T7  
BACKWARD: T3  
Insert length: 847 Std Error: 5.00  
Plate: L30-15 row: A column: 5  
Seq primer: T3  
high quality sequence stop: 450.  
Location/Qualifiers  
1. 837  
/organism="Mesembryanthemum crystallinum"  
/db\_xref="taxon:3544"  
/clone\_lib="L30-1401"  
/tissue\_type="leaf, 30 h 0.4M NaCl"  
/dev\_stage="Six week old"  
/note="Vector: Lambda Uni-Zap XR, Bluescript SK-; Site\_1:  
ECORI; Site\_2: XhoI"  
BASE COUNT 235 a 193 c 219 g 190 t  
ORIGIN

Query Match 5.08; Score 47.2; DB 10; Length 837;  
Best Local Similarity 51.1; Pred. No. 0.16;  
Matches 136; Conservative 0; Mismatches 128; Indels 2; Gaps 1;  
QY 453 AAAGAAGGAGAGTAACTCAAGAAATCAAGTCCCAATGGGAAATTCATACAG 512  
DB 316 AGAGAAGGCTGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 375  
QY 513 TGACATTCGAAAAATCTCTGGAGTTCATTCATTCCTGTAAAACTCAGAGCTTTTCT 572  
DB 376 TGGCATTCGATTCGCTTTTGGTAAAGAGCCGATTAATGGGAGGAGGAGGAGGAGG 435  
QY 573 ACTGTAGGAGTCAAGTCAAGTTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 632  
DB 436 GCTGATTCGATTCATCAAGTTCATCAAGGAGGAGGAGGAGGAGGAGGAGGAGG 495  
QY 633 CAAGACTAATTTTAAATAGAGCATTCATCCGATGATCCCAAGAAATTAATTTTCA 690  
DB 496 GACCCAAAGTGGGTTCATCGGTTCGTAAGGAGGAGGAGGAGGAGGAGGAGGAGG 555  
QY 691 AAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 716  
DB 556 AATCGGGAGAGATTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 581

RESULT 8  
BE130330  
LOCUS 489 bp mRNA linear EST 20-FEB-2001  
DEFINITION L48-48473 Ice plant Lambda Uni-Zap XR expression library, 48 hours  
NaCl treatment Mesembryanthemum crystallinum cDNA clone L48-484 5',  
mRNA sequence.  
ACCESSION BE130330  
VERSION BE130330.1 GI:8577693  
KEYWORDS EST.  
SOURCE common iceplant.  
ORGANISM Mesembryanthemum crystallinum

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
AUTHORS Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
TITLE Caryophyllales; Alzooceae; Mesembryanthemum.  
JOURNAL 1 (bases 1 to 489)  
COMMENT Cushman, J.C.  
An expressed sequence tag database for the common ice plant,  
Mesembryanthemum crystallinum  
Unpublished (1997)  
Contact: Cushman JC  
Department of Biochemistry  
University of Nevada  
MS200, Reno, NV 89557-0014, USA  
Tel: 775-784-1918  
Fax: 775-784-1650  
Email: jcushman@unr.edu  
PCR Primers

Db 313 GCTGATTCCTATACACATGCTTCTGAGCAGCAGCGTTTAAAGTATATTGAAAGTAAGGT 372  
 Qy 633 ---CAAGACTAATTTTAAATAGAGATCTACCTCCGATCCCAAGTAATTAATTTGGAGGA 689  
 Db 373 GACCCAAGTGGTTACATGCTCGTTCATACCCGACCGAAGAGTGTCACTTTTGGAGAA 432  
 Qy 690 GAAGTGGGGCAAAATCTCTGAGGCAATTCACA 721  
 Db 433 CAATTGGGGGAAGATTTCGCAAGAGATTTCATA 464

RESULT 4  
 BE033546  
 LOCUS  
 DEFINITION MF03A09 MF Mesembryanthemum crystallinum cDNA 5' similar to  
 ribosome-inactivating protein, mRNA sequence.  
 ACCESSION BE033546  
 VERSION BE033546.1 GI:8328555  
 KEYWORDS EST.  
 SOURCE common iceplant.  
 ORGANISM Mesembryanthemum crystallinum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Caryophyllidae; Caryophyllales; Alzaceae; Mesembryanthemum.  
 REFERENCE 1 (bases 1 to 1038)  
 AUTHORS Bohnert, H.J., Borchert, C., Brazill, S., Brooks, J., Eaton, M., Ferrer,  
 H., Kawasaki, S., McCollough, A., Michalowski, C.B., Palacio, C.,  
 Scara, G., Wheeler, M. and Zepeda, C.R.  
 TITLE Functional Genomics of Plant Stress Tolerance  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Michalowski, C.B.  
 University of Arizona  
 Bio Sciences West room 513, Tucson, AZ 85721, USA  
 Tel: 520-621-7982  
 Fax: 520-621-1697  
 Email: cbm@u.arizona.edu  
 An open reading frame exists.

FEATURES  
 source  
 1..1038  
 /organism="Mesembryanthemum crystallinum"  
 /db\_xref="taxon:3544"  
 /clone\_lib="MF"  
 /tissue\_type="Root"  
 /dev\_stage="5-6 weeks old"  
 /note="vector: Bluescript SK+; Site\_1: EcoRI; Site\_2:  
 XhoI"

BASE COUNT 232 a 259 c 183 g 363 t 1 others  
 ORIGIN  
 Query Match 5.9%; Score 55.4; DB 10; Length 1038;  
 Best Local Similarity 52.7%; Pred. No. 0.0013;  
 Matches 144; Conservative 0; Mismatches 126; Indels 3; Gaps 1;

Qy 453 AAGAAAGCAGCAAGTAACCAAGCAAGTCAATTCGAATTCGAATTCGAATTCAGCAG 512  
 Db 286 AGAAGAGCGTGCAGCAAGCAAGCGCAATTCGAATTCGAATTCGAATTCAGCAG 345  
 Qy 513 TGCAATTCGAAATATCTCTGGAGTTGATTCATTCCTCTGTAATTCAGCAGCGCTTTTCT 572  
 Db 346 CGCATTCGATCGGTTTGGTAGACCCCAATTAATGGCGAGTTGAGGCGCAATTCCT 405  
 Qy 573 ACTGATGACCAATCAATGGTTTCAGAGCGAGCGGATTCGAATTCAGTACATAGAGAACCAAGT 632  
 Db 406 GCTCATTGCTATACAGATGGTTCTGAGCAGCAGCGTTTAAGTTTATTGAAAGTAAGT 465  
 Qy 633 CA---AGACTAATTTTAAATAGAGATCTCTACCTGATCCCAAGTAATTAATTTGGAGGA 689  
 Db 466 GACCCAAGTGGTTACATGCTGTTTCAACACCGCAGCGAAGAGTGTGAGTTCGGTGA 525  
 Qy 690 GAAGTGGGGCAAAATCTCTGAGGCAATTCACA 722  
 Db 526 CAATTGGGTGAACATTTTCCATGAGACTCATAA 558

RESULT 5  
 T24255  
 LOCUS  
 DEFINITION crs1345 lambdaZAPST Ricinus communis cDNA clone pcrs1345, mRNA  
 sequence.  
 ACCESSION T24255  
 VERSION T24255.1 GI:689074  
 KEYWORDS EST.  
 SOURCE castor bean.  
 ORGANISM Ricinus communis  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eurosids 1; Malpighiales; Euphorbiaceae; Ricinus.  
 REFERENCE 1 (bases 1 to 405)  
 AUTHORS Vandell, F.J., Turner, S. and Somerville, C.  
 TITLE Expressed sequence tags from developing castor seeds  
 JOURNAL Plant Physiology, 108, 1141-1150 (1995)  
 COMMENT Contact: Somerville CR  
 Carnegie Institution  
 Carnegie Institution, 290 Panama St, Stanford, CA 94305  
 Tel: 415/253/521  
 Email: crs@andrew.stanford.edu  
 Seq primer: T3.

FEATURES  
 source  
 1..405  
 /organism="Ricinus communis"  
 /strain="Baker 296"  
 /db\_xref="taxon:3988"  
 /clone\_lib="pcrs1345"  
 /note="vector: lambdaZAPST"  
 /note="vector: lambdaZAPST; Site\_1: EcoRI; Site\_2: XhoI;  
 Poly(A)+ RNA was purified from developing stage 111 to  
 stage V (Greenwood & Bewley, Can. J. Bot. 60:1751-1760,  
 1982) endosperm plus embryo of immature castor fruits.  
 CDNA was synthesized and cloned into lambdaZAPST according  
 to the instructions of the manufacturer (Stratagene);  
 synthesis was primed from the poly(A) tail, and cloned  
 directionally into XhoI (3') and EcoRI (5') sites. In few  
 cases, sequence data indicated that this directionality  
 was reversed. Partial cDNA clones predominated."  
 BASE COUNT 118 a 84 c 86 g 110 t 7 others  
 ORIGIN

Query Match 5.5%; Score 52.4; DB 14; Length 405;  
 Best Local Similarity 59.2%; Pred. No. 0.006;  
 Matches 125; Conservative 0; Mismatches 79; Indels 7; Gaps 2;  
 Qy 555 AACTGAGCGTTTCTTCTAGTGGTTCAGCATCAATGTTTCAGAGCAGCGCATTCAA 614  
 Db 82 AACTCTGCGTCTGCTTCTTATGTTTCATCCAAATGATTTCAGAACGACGACATTCGA 141  
 Qy 615 CTACATGAGACACCAAGTCAAGACTAATTTT-----AATGAGCATTCTAGCTGATCC 668  
 Db 142 CTACATTTCAGGAGAGAAATGCGCAGAGATAGGTACACCGNAGATCTGACCAATCC 201  
 Qy 669 CAAAGTAATTAATTTTCGAGGAGAGAGTGGGCAATCTCTGAGGCAATTCACA-ATGCA 727  
 Db 202 TAGCGTAATTACACTTTNGAATAGTTGGGAGAGACTTCCACTGCAATTCAGAGTCA 261  
 Qy 728 AGAATGAGTCTTTTACCCCAATCACTTCAGCT 758  
 Db 262 ACCAAGGCGCTTTCTTCTAGTCCATTCACCT 292

RESULT 6  
 BE037217  
 LOCUS  
 DEFINITION MF18H02 MP Mesembryanthemum crystallinum cDNA 5' similar to  
 ribosome-inactivating protein/antiviral protein, mRNA sequence.  
 ACCESSION BE037217  
 VERSION BE037217.1 GI:8332233



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 2, 2003, 01:08:52 : Search time 1673 seconds  
(without alignments)  
9148.078 Million cell updates/sec

Title: US-09-978-274A-1  
Perfect score: 945  
Sequence: 1 atgaagtgatgtgttagt.....atctattgaaggttctlaa 945

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST.\*

- 1: em\_estba:\*
- 2: em\_esthum:\*
- 3: em\_estin:\*
- 4: em\_estmu:\*
- 5: em\_estov:\*
- 6: em\_estpl:\*
- 7: em\_estro:\*
- 8: em\_hcc:\*
- 9: gb\_est1:\*
- 10: gb\_est2:\*
- 11: gb\_hcc:\*
- 12: gb\_est3:\*
- 13: gb\_est4:\*
- 14: gb\_est5:\*
- 15: em\_estfun:\*
- 16: em\_estom:\*
- 17: gb\_gss:\*
- 18: em\_gss\_hum:\*
- 19: em\_gss\_inv:\*
- 20: em\_gss\_pln:\*
- 21: em\_gss\_vit:\*
- 22: em\_gss\_fun:\*
- 23: em\_gss\_mam:\*
- 24: em\_gss\_mus:\*
- 25: em\_gss\_other:\*
- 26: em\_gss\_pro:\*
- 27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	87	9.2	993	10	BE035039
C 2	80.6	8.5	816	10	BE035038
C 3	65.6	6.9	1033	10	BE036541
C 4	55.4	5.9	1038	10	BE033546
C 5	52.4	5.5	405	14	T24255
C 6	50.8	5.4	639	10	BE037217

7	47.2	5.0	837	10	AW053634
8	46.6	4.9	489	10	BE130330
9	45.4	4.8	267	9	AA856221
10	44.6	4.7	1021	10	BE036639
11	43.6	4.6	915	17	CNS0296A
C 12	42.4	4.5	561	17	AO156308
C 13	42.4	4.5	1101	17	CNS0039G
C 14	40.8	4.3	1001	17	CNS007RE
15	40.4	4.3	557	10	AW230870
16	40.4	4.3	503	17	AO670952
17	40.4	4.3	579	13	RJ365833
18	40.4	4.3	635	10	AW323815
19	40.4	4.3	685	13	BG920190
C 20	40	4.2	700	13	BM634714
C 21	40	4.2	715	13	BM654071
C 22	40	4.2	895	17	CNS015AJ
23	39.6	4.2	444	14	K31564
24	39.6	4.2	870	17	CNS0187R
25	39.6	4.2	1204	17	CNS016E2
26	39.4	4.2	456	12	BG555640
C 27	39.4	4.2	913	17	CNS00CE20
C 28	39.2	4.1	1101	17	CNS000EPC
C 29	39.2	4.1	1101	17	CNS0100X
C 30	39.2	4.1	1146	17	CNS021C2
31	39	4.1	359	10	AW237651
32	39	4.1	684	14	BM982928
C 33	39	4.1	1019	17	CNS006FW
34	38.8	4.1	587	10	BE533685
C 35	38.6	4.1	593	13	BM610967
C 36	38.6	4.1	721	13	DM574953
37	38.4	4.1	860	17	AZ538875
C 38	38.4	4.1	1101	17	CNS017DU
C 39	38.2	4.0	512	13	AW199192
C 40	38.2	4.0	512	13	BI451206
C 41	38.2	4.0	680	13	RJ078909
42	38.2	4.0	870	17	CNS060U22
43	38	4.0	608	13	RJ397293
44	37.8	4.0	447	9	AA964758
C 45	37.8	4.0	676	13	IM585966

## ALIGNMENTS

RESULT 1  
BE035039/c  
LOCUS BE035039 993 bp mRNA linear EST 07-JUN-2000  
DEFINITION MM02A03 HM Mesembryanthemum crystallinum cDNA 5' similar to  
antiviral protein, mRNA sequence.

ACCESSION BE035039  
VERSION BE035039.1  
KEYWORDS EST  
SOURCE Common iceplant.  
ORGANISM Mesembryanthemum crystallinum

REFERENCE  
AUTHORS Bolner H.J., Rorchert C., Brazille S., Brooks J., Eaton M., Furra H., Kuvusaki S., McCollough A., Michalowski C.U., Palencia C., Scara G., Wheeler M. and Zepeda C.R.  
TITLE Functional Genomics of Plant Stress Tolerance  
J (bases 1 to 993)

COMMENT  
Unpublished (2000)  
Contact: Michalowski, C.B.  
University of Arizona  
Bio Sciences West Room 513, Tucson, AZ 85721, USA  
Tel: 520-621-7982  
Fax: 520-621-7982

FEATURES  
Source  
Insert Length: 1 Std Error: 0.00.  
Location/Qualifiers  
/organism="Mesembryanthemum crystallinum"



241 GCAACCTAAAAACCACTTACACTAATGCTGAGACGAAATAAATTAACCTTATACGTGATGGCTAT 300  
|||||  
241 GCAACCTAAAAACCACTTACACTAATGCTGAGACGAAATAAATTAACCTTATACGTGATGGCTAT 300  
301 TCTGATCCCTTCAATGCGAATAAGTGTGCTTACCATATATTTAATGATATTAACAAGCACC 360  
|||||  
301 TCTGATCCCTTCAATGCGAATAAGTGTGCTTACCATATATTTAATGATATTAACAAGCACC 360  
361 GAACGCACTGATGGGAGAACTCTTTGCTCAAGTTCTAGTTCTCGTGTGCAATGTCC 420  
|||||  
361 GAACGCACTGATGGGAGAACTCTTTGCTCAAGTTCTAGTTCTCGTGTGCAATGTCC 420  
421 ATTAACCTAATAGCTTATATCGACCATGGAAGAAAGCAAGTAACTCAAGAAAT 480  
|||||  
421 ATTAACCTAATAGCTTATATCGACCATGGAAGAAAGCAAGTAACTCAAGAAAT 480  
481 CAAGTCCAATTGGGAATTCAAATCTCAGCAGTCAATTCGAAATAATCTCTGGAGTTGAT 540  
|||||  
481 CAAGTCCAATTGGGAATTCAAATCTCAGCAGTCAATTCGAAATAATCTCTGGAGTTGAT 540  
541 TCATCCCTGTAATAACTAGAGCTTTTCTTCTACTGGTACCCAAATCTCTGGAGTTGAT 600  
|||||  
541 TCATCCCTGTAATAACTAGAGCTTTTCTTCTACTGGTACCCAAATCTCTGGAGTTGAT 600  
601 GCAGCGGATTCAGTACATAGAGCAACCAAGTCAAGACTAATTTTATAGAGCAATCTAC 660  
|||||  
601 GCAGCGGATTCAGTACATAGAGCAACCAAGTCAAGACTAATTTTATAGAGCAATCTAC 660  
661 CTGATCCCAAACTAATTAATTTGGAGGAGAAAGTGGGCAAAATCTCTGAGCAATTCAC 720  
|||||  
661 CTGATCCCAAACTAATTAATTTGGAGGAGAAAGTGGGCAAAATCTCTGAGCAATTCAC 720  
721 AATGCCAAGAAATGGGCTTTTACCACCAACTTGAAGTGGATGCCAAGTACCAAG 780  
|||||  
721 AATGCCAAGAAATGGGCTTTTACCACCAACTTGAAGTGGATGCCAAGTACCAAG 780  
781 TGGATAGTTCTTAGAGTGAATGAATCAATCGTGATGGCACTCTTAAAGTACGTTAAT 840  
|||||  
781 TGGATAGTTCTTAGAGTGAATGAATCAATCGTGATGGCACTCTTAAAGTACGTTAAT 840  
841 GGAACCTGTGAGCACTTACCAAAATGCCATGTTCTCAAGTTATTAATTTCTACTTAT 900  
|||||  
841 GGAACCTGTGAGCACTTACCAAAATGCCATGTTCTCAAGTTATTAATTTCTACTTAT 900  
901 TATAATATATGCTAATCTGTGATCTATTTGAAGGATTCTAA 945  
|||||  
901 TATAATATATGCTAATCTGTGATCTATTTGAAGGATTCTAA 945

RESULT 2  
US-09-978-274A-3  
: Sequence 3, Application US/09978274A  
: Patent No. US20020116737A1  
GENERAL INFORMATION:  
: APPLICANT: Thomas, Christopher  
: APPLICANT: McPherson, Michael  
: APPLICANT: Atkinson, Howard  
: APPLICANT: Neelam, Anil  
TITLE OF INVENTION: PLANT CELL DEATH SYSTEM  
FILE REFERENCE: 9341-028  
CURRENT APPLICATION NUMBER: US/09/978.274A  
PRIOR FILING DATE: 2001-10-15  
PRIOR APPLICATION NUMBER: 0025225.4  
PRIOR FILING DATE: 2000-10-14  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 3  
LENGTH: 792  
TYPE: DNA  
ORGANISM: Phytolacca americana  
US-09-978-274A-3

Query Match 83.2% Score 786.4; DB 10; Length 792;  
Best Local Similarity 99.9%; Pred. No. 3.4e-214;  
Matches 787; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 73 ATAAATACGATCAGCTTTGATGCTGAAATGACCAATTAACAAATATGCAATCTTTATG 132  
|||||  
DB 4 ATAAATACGATCAGCTTTGATGCTGAAATGACCAATTAACAAATATGCAATCTTTATG 63  
|||||  
QY 133 GAATCTCTTCTGTAATCAAGCAAGATCCAAACTAATAATGCTATGGCAATCAATGCTA 192  
|||||  
DB 64 GAATCTCTTCTGTAATCAAGCAAGATCCAAACTAATAATGCTATGGCAATCAATGCTA 123  
|||||  
QY 193 CCGTATACCTAATTCGACCCCTTAAGTACTTATTGCTTAAGCTCCAAAGTCCAAACTTAAA 252  
|||||  
DB 124 CCGTATACCTAATTCGACCCCTTAAGTACTTATTGCTTAAGCTCCAAAGTCCAAACTTAAA 183  
|||||  
QY 253 ACCATTACACTAATGCTGAGAGCAATAAATCTTATAGTGTGCTGCTATTTGCTGATGCTTTC 312  
|||||  
DB 184 ACCATTACACTAATGCTGAGAGCAATAAATCTTATAGTGTGCTGCTATTTGCTGATGCTTTC 243  
|||||  
QY 313 AATGGCAATAAGTGTCTTACCATATATTTAATGATATTACAGCACCGCAAGCTGAT 372  
|||||  
DB 244 AATGGCAATAAGTGTCTTACCATATATTTAATGATATTACAGCACCGCAAGCTGAT 303  
|||||  
QY 373 GTGGCAATACCTTTTCTGCTCAAGTTCTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 432  
|||||  
DB 304 GTGGCAATACCTTTTCTGCTCAAGTTCTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 363  
|||||  
QY 433 AGCTTATATCCGACCATGGAAGAAAGCAAGTAACTCAAGAAATCAATCTCAATTTG 492  
|||||  
DB 364 AGCTTATATCCGACCATGGAAGAAAGCAAGTAACTCAAGAAATCAATCTCAATTTG 423  
|||||  
QY 493 GGAATTCAAATCTCTAGCAGTGCATTTGGAATAATCTCTGGAGTTGATTTCTGCTGTA 552  
|||||  
DB 424 GGAATTCAAATCTCTAGCAGTGCATTTGGAATAATCTCTGGAGTTGATTTCTGCTGTA 483  
|||||  
QY 553 AAAATCAGGCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 612  
|||||  
DB 484 AAAATCAGGCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 543  
|||||  
QY 613 AAGTACATAGCAACCAATCTCAAGACTAATTTTATAGAGCAATTTTACCTGATGCTGCTGCTGCT 672  
|||||  
DB 544 AAGTACATAGCAACCAATCTCAAGACTAATTTTATAGAGCAATTTTACCTGATGCTGCTGCTGCT 603  
|||||  
QY 673 GTAAATTAATTTGAGGAGAACTGGGCAAAATCTCTGAGGCAATTCACAAATGCAAGTAT 732  
|||||  
DB 604 GTAAATTAATTTGAGGAGAACTGGGCAAAATCTCTGAGGCAATTCACAAATGCAAGTAT 663  
|||||  
QY 733 GGGCTTTTACCCAAACCACTTGAAGTGGATGGCAAGTACCAAGTGGATGCTGCTGCTGCTGCT 792  
|||||  
DB 664 GGGCTTTTACCCAAACCACTTGAAGTGGATGGCAAGTACCAAGTGGATGCTGCTGCTGCTGCT 723  
|||||  
QY 793 AGAGTGGATGAATCAATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 852  
|||||  
DB 724 AGAGTGGATGAATCAATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 783  
|||||  
QY 853 ACAACTTA 860  
|||||  
DB 784 ACAACTTA 791

RESULT 3  
US-09-978-274A-19  
: Sequence 19, Application US/09978274A  
: Patent No. US20020116737A1  
GENERAL INFORMATION:  
: APPLICANT: Thomas, Christopher  
: APPLICANT: McPherson, Michael  
: APPLICANT: Atkinson, Howard  
: APPLICANT: Neelam, Anil  
TITLE OF INVENTION: PLANT CELL DEATH SYSTEM  
FILE REFERENCE: 9341-028  
CURRENT APPLICATION NUMBER: US/09/978.274A

us-09-978-274a-1.rnpb

Mon Jun 2 10:54:58 2003

Patent No. US20020116737A1  
GENERAL INFORMATION:  
APPLICANT: Thomas, Christopher  
APPLICANT: McPherson, Michael  
APPLICANT: Alkinson, Howard  
APPLICANT: Neelam, Anil  
TITLE OF INVENTION: PLANT CELL DEATH SYSTEM  
FILE REFERENCE: 9341-028  
CURRENT APPLICATION NUMBER: US/09/978-274A  
CURRENT FILING DATE: 2001-10-15  
PRIOR APPLICATION NUMBER: 0025225.4  
PRIOR FILING DATE: 2000-10-14  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 30  
LENGTH: 1375  
TYPE: DNA  
ORGANISM: Phytolacca americana  
US-09-978-274A-30

Query Match 80.6%; Score 761.4; DB 10; Length 1092;  
Best Local Similarity 99.9%; Pred. No. 5.7e+207;  
Matches 762; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
73 ATAAATACGATCAGCTTTGATGCTGGAATGCCACCATTAACAAATATGCCACCTTTATG 132  
4 ATAAATACGATCAGCTTTGATGCTGGAATGCCACCATTAACAAATATGCCACCTTTATG 63  
133 GAATCTCTTCGTAATCAAGCGAAGATCCAAACTAAATGCTATGGCATACCAATGCTA 192  
64 GAATCTCTTCGTAATCAAGCGAAGATCCAAACTAAATGCTATGGCATACCAATGCTA 123  
193 CCGTACTAATTCGACCCCTAAGTACTTATTGGTTAAGCTCCAAAGTGCAACCTTAA 252  
124 CCGTACTAATTCGACCCCTAAGTACTTATTGGTTAAGCTCCAAAGTGCAACCTTAA 183  
253 ACCATTACACTAATCTGAGAGCAATAACTTATACGTTGATGGGCTATTCTGATCCCTTC 312  
184 ACCATTACACTAATCTGAGAGCAATAACTTATACGTTGATGGGCTATTCTGATCCCTTC 243  
313 AATGGCAATTAAGTGTGCTTACCATATATTATGATATTACAGAGCGAAGCACTGAT 372  
244 AATGGCAATTAAGTGTGCTTACCATATATTATGATATTACAGAGCGAAGCACTGAT 303  
373 GTGAGNACTCTTCTGCTCAAGTCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 432  
304 GTGAGNACTCTTCTGCTCAAGTCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 363  
433 AGCTTATATCCGACCATGGAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 492  
364 AGCTTATATCCGACCATGGAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 423  
493 GGAATTAATATCTACGAGTGCATGGAAGAAATCTCTGGAGTTGATTTCCCTGTA 552  
424 GGAATTAATATCTACGAGTGCATGGAAGAAATCTCTGGAGTTGATTTCCCTGTA 483  
553 AAAACTGAGGCTTTTCTTCTACTGCTAGCCATCCAAATGTTTTCAGAGCGAGCGGATTC 612  
484 AAAACTGAGGCTTTTCTTCTACTGCTAGCCATCCAAATGTTTTCAGAGCGAGCGGATTC 543  
613 AAGTACATAGAGAACCAAGTCAAGACTAATTTTATAGAGCTTCTACCTGATCCCAAA 672  
544 AAGTACATAGAGAACCAAGTCAAGACTAATTTTATAGAGCTTCTACCTGATCCCAAA 603  
673 GTAATTAATTTGGAGAGAGTGGGCAAAATCTCTGAGGCAATTCACAAATGCCAAGAT 732  
604 GTAATTAATTTGGAGAGAGTGGGCAAAATCTCTGAGGCAATTCACAAATGCCAAGAT 663  
733 GGGGCTTTTACCAAAACCACTTGAAGTGTGATGGGCAAAAGGTACCAAGTGGATAGTTCT 792  
664 GGGGCTTTTACCAAAACCACTTGAAGTGTGATGGGCAAAAGGTACCAAGTGGATAGTTCT 723  
793 AGAGTGGATGAATCAATCTGATGTTGGGCACTCTCTTAAGTACG 835  
724 AGAGTGGATGAATCAATCTGATGTTGGGCACTCTCTTAAGTACG 766

Query Match 66.9%; Score 632.4; DB 10; Length 1379;  
Best Local Similarity 80.0%; Pred. No. 4.3e-170;  
Matches 757; Conservative 0; Mismatches 186; Indels 3; Gaps 1;  
3 GAAGCTGATGCTTGTAGTGTGGTGACGTTTAAAGCTGGCTCATTGCTGCACAAATTC 62  
221 GAAGATGAAGTGCATGCTGTGTCGACAAATATAATATGGCTCATCTTGCACATTC 280  
63 AACTTGTGCTCAATATACATACATCTTTCATGCTGGAATGCCACCATTAACAAATATGC 122  
281 AACTTGTGCTGTAATACATCAATCACTTACAAATGTGGAGTCCACCATTAAGCAATATGC 340  
123 CACCTTTATGGAATCTCTGCTAATCAAGCAAGATCCAAACTAAATATGCTATGCAAT 182  
341 CACCTTTTGAATGATCTTCTGCTAATCAAGCAAGATCCAAAGTTTAAATGCTATGCAAT 400  
183 ACCAATGCTACCTGATACATTAATTCGACCCCTAAGTACTTATGCTTAAAGCTTCAAGCT 242  
401 ACCAATGCTCCCAATACAAATACAAATCCAAAGTACGTTGCTGGAGCTCAAGCTTTC 460  
243 AAGCTTAAACCATTAACATCACTAATGCTGAGCAAGAAATACCTTATAGCTATGCTATTC 302  
461 AATATAAAACCATCACTAATGCTGAGCAAGAAATACCTTATGCTATGCTATGCTATTC 520  
303 TGATCCCTTCAATGCAATTAAGTGTCTTACCATATATTTAATGATATTAACAAATGCA 362  
521 TGATCCCTTGAACCAATTAAGTGTCTTACCATATCTTTAATGATATTCAGGCTATTC 580  
363 AGCACTGATGTGGAGAACTCTTGTCTCAAGTCTTCTGCTGCTGCTGCTGCTGCTGCT 422  
581 AGCACTGATGTGGAGAACTCTTGTCTCAAGTCTTCTGCTGCTGCTGCTGCTGCTGCT 640  
423 TAATCACTAATGCTTATATCCGACCATGGAAGAAAGAGAGAGAGAGAGAGAGAGAG 482  
641 AACTTGTGCTCAATGATCCCAATTCACATTCGCAATCAAGCGGAGTAAATCAAGAACTCA 700  
483 ACTCCAAATGGCAATTCAAATACCTGAGCAAGTGCATTCGGAATATCTCTGAGTATTC 542  
701 GGTCAACTGCGAATTCAAATACCTGAGCAAGTGCATTCGGAATATCTCTGAGTATTC 760  
543 ATTCTCTTAAAACTGAGGCTTTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 602  
761 ATTCACTGAGAAACCGAGCGGATTCCTTATTTGGTAGCATCAATATGGATGCTGAGG 820  
603 AGCGGATCACTACATAGAGAACTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 662  
821 AGCAAGATTCAGTACATAGAGAACTCAGGTGAGAACTAATTTTACAGAGCAATTCAG 880  
663 TGATCCCAAGTAAATTAATTTGAGGAGAGAGTGGGCAAAATCTCTGAGGCAATTCAG 722  
881 TAATCCCAAGTAACTTAATTTGAGGAGAGAGTGGGCTAAGTATTCAGCAAGCAATTCAG 940  
723 TGCCAAAGATGGGCTTTTACCAAAACCACTTGAAGTGTGATGCTGAGTGGCAAGTCAAG 782



941 TCCCAAGATGGAGTTTACCCAAACCTCTCGAGTAGTGGATGCCAGTGGTCCCAAGTG 1000  
783 GATAGTCTTAGAGTGGATGAATCAATCTGATGGGCACTCTTAAAGTACCTTAATGG 842  
1001 GATAGTCTTAGAGTGGATGAATCAATCTGATGGGCACTCTTAAAGTACCTTAATGG 1060  
843 AACCTCTGACAGCAACTT---ACCAAAATGCCATGTTCTCTCAAGTTATATTTCTACTTA 899  
1061 GACCTGTCAGCAACTTATAACCAAAATGCCATGTTCTCTCAAGTTATATTTCTACTTA 1120  
900 TTATAATTATATCTCTAATCTTGGTGATCTATTGGAAGGATCTCAA 945  
1121 TTATAATTATATCTCTAATCTTGGTGATCTATTGGAAGGATCTCAA 1166

## RESULT 5

US-09-978-274A-31

; Sequence 31, Application US/09978274A

; Patent No. US20020116737A1

; GENERAL INFORMATION:

; APPLICANT: Thomas, Christopher

; APPLICANT: McPherson, Michael

; APPLICANT: Atkinson, Howard

; TITLE OF INVENTION: PLANT CELL DEATH SYSTEM

; FILE REFERENCE: 9341-028

; CURRENT APPLICATION NUMBER: US/09/978,274A

; PRIOR FILING DATE: 2001-10-15

; PRIOR FILING DATE: 2000-10-14

; NUMBER OF SEQ ID NOS: 32

; SOFTWARE: Patent in version 3.1

; SEQ ID NO 31

; LENGTH: 1368

; TYPE: DNA

; ORGANISM: Phytolacca americana

US-09-978-274A-31

Query Match 66.6%; Score 629.2; DB 10; Length 1368;  
Best Local Similarity 79.8%; Pred. No. 3.5e-169;  
Matches 755; Conservative 0; Mismatches 188; Indels 3; Gaps 1;  
3 GAAGTGTGCTGTGATGTTGGTGACGTTAATAGCGTGGCTCATCTGTCACCAACTTC 62  
221 GAAGTGTGCTGTGATGTTGGTGACGTTAATAGCGTGGCTCATCTGTCACCAACTTC 280  
63 AACTTGTGCTGTGATGTTGGTGACGTTAATAGCGTGGCTCATCTGTCACCAACTTC 122  
281 AACTTGTGCTGTGATGTTGGTGACGTTAATAGCGTGGCTCATCTGTCACCAACTTC 340  
123 CACCTTGTGCTGTGATGTTGGTGACGTTAATAGCGTGGCTCATCTGTCACCAACTTC 182  
341 CACCTTGTGCTGTGATGTTGGTGACGTTAATAGCGTGGCTCATCTGTCACCAACTTC 400  
183 ACCAATGTGCTGTGATGTTGGTGACGTTAATAGCGTGGCTCATCTGTCACCAACTTC 242  
401 ACCAATGTGCTGTGATGTTGGTGACGTTAATAGCGTGGCTCATCTGTCACCAACTTC 460  
243 AACCTTGTGCTGTGATGTTGGTGACGTTAATAGCGTGGCTCATCTGTCACCAACTTC 302  
461 AACCTTGTGCTGTGATGTTGGTGACGTTAATAGCGTGGCTCATCTGTCACCAACTTC 520  
303 TGATCCCTTCAATGCTGTGATGTTGGTGACGTTAATAGCGTGGCTCATCTGTCACCAACTTC 362  
521 TGATCCCTTCAATGCTGTGATGTTGGTGACGTTAATAGCGTGGCTCATCTGTCACCAACTTC 580  
363 AGCGACTGTGCTGTGATGTTGGTGACGTTAATAGCGTGGCTCATCTGTCACCAACTTC 422  
581 AGCGACTGTGCTGTGATGTTGGTGACGTTAATAGCGTGGCTCATCTGTCACCAACTTC 640  
423 TAACCTGTGCTGTGATGTTGGTGACGTTAATAGCGTGGCTCATCTGTCACCAACTTC 482

## RESULT 6

US-09-978-274A-5

; Sequence 5, Application US/09978274A

; Patent No. US20020116737A1

; GENERAL INFORMATION:

; APPLICANT: Thomas, Christopher

; APPLICANT: McPherson, Michael

; APPLICANT: Atkinson, Howard

; APPLICANT: Neelam, Anil

; TITLE OF INVENTION: PLANT CELL DEATH SYSTEM

; FILE REFERENCE: 9341-028

; CURRENT APPLICATION NUMBER: US/09/978,274A

; PRIOR FILING DATE: 2001-10-15

; PRIOR FILING DATE: 2000-10-14

; NUMBER OF SEQ ID NOS: 32

; SOFTWARE: Patent in version 3.1

; SEQ ID NO 5

; LENGTH: 465

; TYPE: DNA

; ORGANISM: Phytolacca americana

US-09-978-274A-5

Query Match 48.6%; Score 459; DB 10; Length 465;

Best Local Similarity 100.0%; Pred. No. 7.2e-121;

Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

73 ATAATAGCATCACCCTTTGATGCTGGAATGCCACCATTAACAATATGCCACTTTTANG 132  
4 ATAAATAGCATCACCCTTTGATGCTGGAATGCCACCATTAACAATATGCCACTTTTANG 63  
133 GAATCTCTTCTGATCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 192  
64 GAATCTCTTCTGATCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 124  
143 CTTGATCAATATGAGGCTTAACTTATTTAGGCTGCAAGCTGCAAGCTGCAAGCTGCAAG 252  
124 CTTGATCAATATGAGGCTTAACTTATTTAGGCTGCAAGCTGCAAGCTGCAAGCTGCAAG 184



640 GATGCCAATGGGTATCAGCCAGATCTTAAAGCTATTTCCTAGAGAAAATTTGGCAGCT. 699  
 703 ATCTCTGAGGCAATTCACAATGCCAAGATGGGGCTTTACCCCAACCACTTTGAGCTAGTG 762  
 700 GTTCTAAGGCAATTCGCAAAAGTTGGCACCCTCCGGTGATAGTACTGTTACTTTACCTGGA 759  
 763 GATGCCAAGGTACCAAGTGGATAGTCTTCTAGAGTGGATGAATCAATCGTGTGGA 822  
 760 GACCTAAAGATGACAATAATAAACCCTTGGACTAGGGCACCACCAAGCACCCTTAAGAAC 819  
 823 CTCCTTAAGTAGGTTAAATGGAACCTGTCAGACAACCTTACCAAAATGCCATGTTCTCTCAA 882  
 820 GACATATGGCACTCCTCAACCCACCTTACTTCCAAAGTTAAAGTTCCATGTTCCCTGAA 879  
 883 GTTATAATTTCTACTATTATATAATATATATGCTTAATCTTGGTGA 926  
 880 ATTATGCTCTATTATTATAGGACTAGTATTAGTAACCTTGGTGA 923

## RESULT 9

US-09-861-257-37  
 : Sequence 37, Application US/09861257  
 : Publication No. US20030040496A1

## GENERAL INFORMATION:

: APPLICANT: Chandier, Lois Ann  
 : APPLICANT: Sosnowski, Barbara A.  
 : APPLICANT: Baird, J. Andrew  
 : APPLICANT: Pierce, Glenn

: TITLE OF INVENTION: TREATMENT OF TUMORS USING

: TITLE OF INVENTION: LIGAND-DIRECTED NUCLEIC ACID DELIVERY VEHICLES

: NUMBER OF SEQUENCES: 103

: CORRESPONDENCE ADDRESS:

: ADDRESS: Seed Intellectual Property Law Group

: STREET: 701 Fifth Avenue, Suite 6300

: CITY: Seattle

: STATE: Washington

: COUNTRY: USA

: ZIP: 98104

: COMPUTER READABLE FORM:

: MEDIUM TYPE: Floppy disk

: COMPUTER: IBM PC compatible

: OPERATING SYSTEM: PC-DOS/MS-DOS

: SOFTWARE: PatentIn Release #1.0, Version #1.30

: CURRENT APPLICATION DATA:

: APPLICATION NUMBER: US/09/861,257

: FILING DATE: 17-MAY-2001

: CLASSIFICATION:

: ATTORNEY/AGENT INFORMATION:

: NAME: Christiansen Ph.D., William T.

: REGISTRATION NUMBER: 44,614

: REFERENCE/DOCKET NUMBER: 760100.423C1

: TELECOMMUNICATION INFORMATION:

: TELEPHONE: (206) 622-4900

: TELEFAX: (206) 682-6031

: INFORMATION FOR SEQ ID NO: 37:

: SEQUENCE CHARACTERISTICS:

: LENGTH: 804 base pairs

: TYPE: nucleic acid

: STRANDEDNESS: double

: TOPOLOGY: unknown

: MOLECULE TYPE: cDNA

: FEATURE:

: NAME/KEY: CDS

: LOCATION: 1..804

: FEATURE:

: NAME/KEY: misc-feature

: LOCATION: 1..804

: OTHER INFORMATION: /note= "Nucleotide sequence

: OTHER INFORMATION: corresponding to the clone M13 mp18-G7"

: FEATURE:

: NAME/KEY: mat-peptide

: LOCATION: 46..804

: OTHER INFORMATION: /product= "Saporin"

: Query Match

: Best Local Similarity 9.4%; Score 88.4; DB 9; Length 804;

: Matches 376; Conservative 49.3%; Pred. No. 6.1e-15;

: Mismatches 0; Mismatches 371; Indels 15; Gaps 5;

OY 47 TTGCTGACCAACTTCAACTTGTGCTGCTTAATAATAGATCACCCTTTGTGCTGGAATATGCA 106  
 DB 20 TTTCAGCTTGGCAACCACTGATCGGTCACATCAATCATTAGATCTAATAATATGCA 79  
 OY 107 CCATTAAACAATATGCCACCTTTATGGAATCTCTTGGTAATCAAGCGAAGATATCAAAAC 166  
 DB 80 CCGCGGGTCAATATCTATCTCTTTGTGATTAATTCGAAACAAGTAAAGTATCAAAAC 139  
 OY 167 TAAATGCTATGGCATACCAATGC---TACCTGATATCTAATTCGACCCCTTAAGTATCTAT 223  
 DB 140 TCAATACGCTGGTACCGACATAGCCGCTGATAGCCCACTTCTTAAGAAATATCTCTTA 199  
 OY 224 TGGTTAAGCTCCCAAGGTGCAAAACCTTAATAACCATTAACATCTGTAGAGCAATTAAC 283  
 DB 200 GATTTAATTTCCAAAGTTC---CGAGGACGGTCTCCTTGGCCTAATAAGTATTAAC 256  
 OY 284 TATAGCTGATGGGCTATTTCTGATCCCTTCAATGGCAATAAGTCTGTTTACATATATTTA 343  
 DB 257 TGTATGTTGGTCCGCTATCTTCCATGGATACACGAATGTTAATCGGCGCATATTTCTTCA 316  
 OY 344 ATGATATTACAGCACCCGACGCACTGATGGAGATATCTCTTTGCTCAAGTCTTACTTT 403  
 DB 317 GATCAGAAATTAATCTTCCGCGGAGTTAATCCGCCCTTTTCCAGAGGCGCAATTCGAATC 376  
 OY 404 CTCGTGTTGCAATGTCATTAACTACATACTTATATCCGACCATGCAAAAGAAACGAG 463  
 DB 377 AGAAGCTTTAGATACACACAGATTAATCAGTATCGATCGAAGAAAGATCCCGGATACAC 436  
 OY 464 AAGTA---AATCAAGAAATCAAGTCCCAATTTGGAATTTCAATATCTCAGCAGTCAATG 520  
 DB 437 AGGGAGATTAATCAAGAAAGAACTCGGTTGGGATCGGACTTACTTTTGAAGTCCATGG 496  
 OY 521 GAAATCTCTGGAGTGTGATTCATCTCCCTGTAAAACTGAGGCTTTTTCCTACTGTAG 580  
 DB 497 AGCAGTGAACAAGA---AGGACGTGTGTTAAAGGAGGATAGATTCCTTCTTATGG 553  
 OY 581 CCATCCAAATGTTTTCAGAGGCGGCGGATTCAGATACATAGAGAAACAATTAAGACTA 640  
 DB 554 CTATTCAGATCAAGGCTGAGGCGGCGGATTTAGGTACATACAAACTTCTTAATCAAGA 613  
 OY 641 ATTTTAATAGAGCATTTTACCTGATCCCAAGTAAATTAATTTGGAGGCAAGTGAAGCA 700  
 DB 614 ACTTTCGCAACAATTTCACTCGGAAACAAGTATTCATTTCAAGTTAAATTTGAAAA 673  
 OY 701 AAATCTCTGAGGCAATTCAG---AATGCAAGATGGGCTTTACCCCAACCACTTCAGC 757  
 DB 674 AAATTTCTACGGCAATATACGGGATGCGCAAAACGGGCTTTTAATAAATATATGATT 733  
 OY 758 TAGTGGATGCCAAAGTACCAAGTGTATAGTCTTTAGAGTGG 799  
 DB 734 TCGGCTTTGGAAAGTCAAGGCGGCGGCTTCCAAATCG 775

## RESULT 10

US-09-861-257-38

: Sequence 38, Application US/09861257

: Publication No. US20030040496A1

: GENERAL INFORMATION:

: APPLICANT: Chandier, Lois Ann

: APPLICANT: Sosnowski, Barbara A.

: APPLICANT: Baird, J. Andrew

: APPLICANT: Pierce, Glenn

: TITLE OF INVENTION: TREATMENT OF TUMORS USING

: TITLE OF INVENTION: LIGAND-DIRECTED NUCLEIC ACID DELIVERY VEHICLES

: NUMBER OF SEQUENCES: 103

: CORRESPONDENCE ADDRESS:

us-09-978-274a-1.rnbp

Non Jun 2 10:54:58 2003

ADDRESSEE: Seed Intellectual Property Law Group  
STREET: 701 Fifth Avenue, Suite 6300  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104

COMPUTER: IBM PC compatible  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release 41.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/861,257  
FILING DATE: 17-MAY-2001

CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Christiansen Ph.D., William T.  
REGISTRATION NUMBER: 44,614  
REFERENCE/DOCKET NUMBER: 760100.423C1

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 622-6031

INFORMATION FOR SEQ ID NO: 38:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 804 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: unknown  
MOLECULE TYPE: cDNA

FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..804  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 1..804

OTHER INFORMATION: /note= "Nucleotide sequence  
corresponding to the clone M13 mpl8-G9"

OTHER INFORMATION: /product= "Saporin"  
US-09-861-257-38

Query Match 9.4%; Score 88.4; DB 9; Length 804;  
Best Local Similarity 49.3%; Pred. No. 6.1e-15;  
Matches 376; Conservative 0; Mismatches 371; Indels 15; Gaps 5;

47 TTGCTGACCACTTCAACTGTGGCCATAAATACGATCACCCTTTGATGCTGGAAATGCCA 106  
20 TTTCAGCTGGCAACACTGTGGCCATCAATCAATCATTAGATCTAGTAAATCCGA 79

107 CCATTAACAATAATGCCACCTTTATGGAATCTCTTGTATCAAGCGAAGATCCAAAC 166  
80 CCGCGGTCAATCTACTCTTTTGTGATAAATCCGAACAACGTAAGATGCCAAACC 139

167 TAAATGCTATGGCATACCAATGC---TACTGTATAATTCGACCCCTTAAGTACTTAT 223  
140 TGAATACGGTGTGACGATAGCGGTGATAGCCCACTCTTAAGAGAAAATTCCTTA 199

224 TGGTTAGCTCCAGGTGCAACCTAAACCAATTAATCAATGCTGAGACGAAATCACT 283  
200 GAATTAATTTCCAAAGTTC---CCGAGGACGGTCTCACTTGGCCATAAACCGCATACT 256

284 TATACGCTAGGGGTATCTGATCCCTTCAATGCGAATAAGTCTGTTACCATATATTA 343  
257 TGTATGTCGCGTATCTTGCATGGAATAACCAATGTTAATCGGGCATATTAATTTCA 316

344 ATGATATTACAGCAACCAACCGCACTGATGTGGAGAACTACTCTTTGCTCAAGTTCTAGT 403  
317 CATCAGAAATTAATCTCCCGCGAGTTAACCGCCCTTTTCCAGAGCCCACTGCAATC 376

404 CTCGCTGTCATGTCATTAATCACTACATAGCTTATATCCGACCATGGAAGAAAGCAG 463

Db 377 AGAAGCTTTAGATACACAGAGATTATCAGTCGATTGAAAGAAATCCCGAGATAACAC 436  
Qy 464 AAGTAAACTCA---AGAAATCAAGTCCCAATTGGGAATTCAAATCACTCAGCAGTCACTTG 520  
Db 437 AAGGAGATCAAAAGTAGAAAAGCAACTCGGCTTGGGATTGACTTACTTTCAACGCTCATGG 496  
Qy 521 GAAAAATCTCTGGAGTTCATTTCATTCCTCTGTAATAAAGTGAAGGCTTTTCTTTTATTTAG 580  
Db 497 AAGCAGTGAACAAGA---AGGCAGCTGTGCTTAAGACGCAAGCTAGATTCTCTTCATG 553  
Qy 581 CCATCCAAATGTTTTCAGAGCGAGCGGATTCAGATACATAGAGAACCAACATCAACACTA 640  
Db 554 CTATTCAGATGACGCGCTGAGCGAGCGGATTTAGGTACATACAAAACTTTGGTAATCAAGA 613  
Qy 641 ATTTAATAGACGATTTCTACCTGATCCCAAGTAAATTAATTTGGAGGAGAAATGAGGATA 700  
Db 614 ACTTCCCAACAAGTTCAACTCCGAAACAAGATGATTCAGTTTGAGCTTAACATGAGAAA 673  
Qy 701 AATCTCTGAGCAATTCAC---AATGCCAAGAAATGGGCTTTTACCCAAACACCTTCAGC 757  
Db 674 AATTTCTAGCGCAATATACGGGATGCCCAAAACGCGGTGTTTAAATAAGATTTATGATT 733  
Qy 758 TACTGGATGCCAAAGCTTACCAATGGAATGATCTCTTACAGTGG 799  
Db 734 TCGGTTTGGAAAGTGAAGCGAGTCAAGGACTTGCAAATGG 775

RESULT 11

US-09-792-793A-66  
Sequence 66, Application US/09792793A  
Patent No. US20020168170A1

GENERAL INFORMATION:  
APPLICANT: McDonald, John R.

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING SECONDARY TISSUE DAMAGE

OTHER INFORMATION: OTHER INFLAMMATORY CONDITIONS AND DISORDERS

FILE REFERENCE: 25020-6010

CURRENT APPLICATION NUMBER: US/09/792,793A

CURRENT FILING DATE: 2001-02-22

NUMBER OF SEQ ID NOS: 93

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 66

LENGTH: 765

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: Construct encoding

OTHER INFORMATION: Methionine-Saporin fusion protein

NAME/KEY: CDS

LOCATION: (21)..(765)

US-09-792-793A-66

Query Match 8.9%; Score 84; DB 9; Length 765;  
Best Local Similarity 47.3%; Pred. No. 1.1e-13;  
Matches 320; Conservative 0; Mismatches 350; Indels 6; Gaps 2;

Qy 74 TAAATAGCATCACCCTTGTGATGCTGGAATGCCACCAATTAACAAATATGCCACTTTATCG 133  
Db 55 TTACTAGTATATACCTGGAGCTGCTCAATCCGACCGCGGCCCAATATAGCAGCTTCTCG 64

134 AATCTCTCTGTAATCAAGCGAAGATGCCAAATCTAAATGCTATGTCATACCAATATCTAC 193  
Qy 65 ATAAGATTCGTACACAGCTAAAGATCCGAATCTGAATACGCTGATGATGATTTGCG 124

194 CTGATACCTAATTCGACCCCTTAAGTACTTATTTGTTAGCTCCAGGTCGAACCTTAAAA 253  
Db 125 TCATCGGTCCCGCGCAAGCAAGAAAGTTCCTGCCATTAATCTTTCAAGCTTCGATGCA 184

254 CCATTACTAATCTCAGACGCAAGTAACCTTATACGTGATGGCGCTATTCGATGCTTCA 313  
Qy 185 CTGTTCTCTGGGGCTCAAGCGCCATACCTCTATGTTGTTGCTATCTGCGCATGATA 244

314 ATGCAATTAAGTCTGCT---TACCATATATTTAATGATATTATACAGCACTGGAACGACTG 370

245 ATACGACGCTGAACCGCCCTACTACTTTCTGAGGAGATTACGAGCGGGAATCACTG 304  
371 ATGTGGAGAATACTCTTTGCTCAAGTCTTAGTCTGTTGCAATGTCCATTAACTACA 430  
305 CTCTGTTCCGGAGGCGACCACTGCAACCAACCAAGAGCACTGCAATATACGGAAGATTACC 364  
431 ATAGCTTATATCCGACCATGGAAGAAGAGAGAGTAAGTAACACTCAAGATCAAGTCCAT 490  
365 ATCCATCAGAGAAGCGGAGATCAACCGGCGGATCAATCCCGCAAGAACTGGGTC 424  
491 TGGGAATTCAAATACACTACAGCACTGACATTTGGAATAATCTCTGGAGTTGATTCTCCCTG 550  
425 TGGGTATTGATCTGTGACGACGAGCATGAGAGCGGTCAACAAA---AAGCTCGGGTG 481  
551 TAAAACTGAGCGCTTTTCTTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 610  
482 TTAAGACGAGCGCGCTTCTGCTGATCGCCATTGAGTGTGAGTGTGAGTGTGAGTGTG 541  
611 TCAAGTACATACAGAACCAAGTCAAGACTAATTTTAATAGAGCAATCTACCTGATCCCA 670  
542 TCCGCTACATTCAGAACCTGTGATCAAAACTTCCGGAACAGTTCAATCCGGAATA 601  
671 AAGTAATTAATTTGGAGGAGAGTGGGCAAAATCTCTGAGGCAATTCACAATGGCAAGA 730  
602 AAGTCATTCAGTTCGAGGTTAATTTGGAATAATTTCCACCGCCATTATGTTGAGCGGA 661  
731 ATGGGGCTTTACCCAA 746  
662 AGAACGGTGTTCATCA 677

RESULT 12  
US-09-792-793A-60  
Sequence 60, Application US/09792793A  
Patent No. US20020168370A1  
GENERAL INFORMATION:  
APPLICANT: McDonald, John R.  
APPLICANT: Coghins, Philip  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING SECONDARY TISSUE DAMAGE AND  
TITLE OF INVENTION: OTHER INFLAMMATORY CONDITIONS AND DISORDERS  
FILE REFERENCE: 25020-601D  
CURRENT APPLICATION NUMBER: US/09/792,793A  
CURRENT FILING DATE: 2001-02-22  
NUMBER OF SEQ ID NOS: 93  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 60  
LENGTH: 984  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Construct encoding chemokine-  
LOCATION: (1)...(984)  
NAME/KEY: CDS  
US-09-792-793A-60

Query Match 8.9% Score 84: DB 9: Length 984:  
Best Local Similarity 47.3% Pred. No. 1.2e-13:  
Matches 320: Conservative 0: Mismatches 350: Indels 6: Gaps 2:  
74 TAAATACCATCACTTGTGATGCTGGAATGCAATGCAATTAACAAATATGCGACCTTTATGG 133  
224 TTACTAGTATTACCTGGACCTGGTCAATCCGACCGCGGCAATATACCACTTCCTGG 283  
134 ATCTCTTCTGTAATCAAGCAAGATCCAAACTAAATGCTATGCGATACCAATGCTAC 193  
284 ATAAGATCGTAACACGCTAAAGATCCGAATCTGAATACGCTGCTACTGATATGGCG 343  
194 CTGATCACTAATTCGACCCCTTAGTACTTATTGGTAAAGTCCCAAGTCCCAACCTTAARA 253  
344 TCATCGTCCGCGCAAGAAAGTTCCTCGCATTAATTTCAAGCTCCCGTGCA 403

254 CCAATTACACTAATGCTGAGAGCAATAAATTATAGGTGATGGCTATTCTCAATCCCTTCA 313  
404 CTGTTTCTCTGGGCTGAAGCGGCGATACCTGTATGTTGCTTCTGCTGCGATGATA 463  
314 ATGSCAATAAGTGTCT---TACCATATATTAAATGATATTACAAGCACTCAATGCACTG 370  
464 ATACGAACCTGAACCGCGCTACTACTTTCTGAGGAGATTACGAGCGGGAATCCACTG 523  
371 ATGTGGAGAAATACTCTTTGCTCAAGTCTTAGTCTCTGCTGTTGCAATGCTCCATTAACTACA 430  
524 CTCCTGTTCCCGAGGCGACCACTGCAACCAACCAAGCACTGCAATATACGGAAGATTACC 583  
431 ATAGCTTATATCCGACCATGGAAGAAGAGAGTAAGTAACACTCAAGATCAAGTCCAT 490  
584 ATCCATCAGAGAAGCGGAGATCAACCGGCGGATCAATCCCGCAAGAACTGGGTC 643  
491 TGGGAATTCAAATACACTACAGCACTGACATTTGGAATAATCTCTGGAGTTGATTCTCCCTG 550  
644 TGGGTATTGATCTGTGACGACGAGCATGAGAGCGGTCAACAAA---AAGCTCGGGTG 700  
551 TAAAACTGAGCGCTTTTCTTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 610  
701 TAAAGACGAGCGCGCTTCTGCTGATCGCCATTGAGTGTGAGTGTGAGTGTGAGTGTG 760  
611 TCAAGTACATACAGAACCAAGTCAAGACTAATTTTAATAGAGCAATCTACCTGATCCCA 670  
761 TCCGCTACATTCAGAACCTGTGATCAAAACTTCCGGAACAGTTCAATTCGGAATA 820  
671 AAGTAATTAATTTGGAGGAGAGTGGGCAAAATCTCTGAGGCAATTCACAATGGCAAGA 730  
821 AAGTCATTCAGTTCGAGGTTAATTTGGAATAATTTCCACCGCCATTATGTTGAGCGGA 880  
731 ATGGGGCTTTACCCAA 746  
881 AGAACGGTGTTCATCA 896

RESULT 13  
US-09-792-793A-63  
Sequence 63, Application US/09792793A  
Patent No. US20020168370A1  
GENERAL INFORMATION:  
APPLICANT: McDonald, John R.  
APPLICANT: Coghins, Philip  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING SECONDARY TISSUE DAMAGE  
TITLE OF INVENTION: OTHER INFLAMMATORY CONDITIONS AND DISORDERS  
FILE REFERENCE: 25020-601D  
CURRENT APPLICATION NUMBER: US/09/792,793A  
CURRENT FILING DATE: 2001-02-22  
NUMBER OF SEQ ID NOS: 93  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 63  
LENGTH: 993  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)...(993)  
OTHER INFORMATION: Description of Artificial Sequence: Construct encoding chemokine  
OTHER INFORMATION: toxin fusion protein BOTA-XIN-AM-SAPORIN  
US-09-792-793A-63

Query Match 8.9% Score 84: DB 9: Length 993:  
Best Local Similarity 47.3% Pred. No. 1.2e-13:  
Matches 320: Conservative 0: Mismatches 350: Indels 6: Gaps 2:  
74 TAAATACCATCACTTGTGATGCTGGAATGCAATGCAATTAACAAATATGCGACCTTTATGG 133  
233 TTACTAGTATTACCTGGACCTGGTCAATCCGACCGCGGCAATATACCACTTCCTGG 292  
134 ATCTCTTCTGTAATCAAGCAAGATCCAAACTAAATGCTATGCGATACCAATGCTAC 193  
292 ATAAGATCGTAACACGCTAAAGATCCGAATCTGAATACGCTGCTACTGATATGGCG 343  
194 CTGATCACTAATTCGACCCCTTAGTACTTATTGGTAAAGTCCCAAGTCCCAACCTTAARA 253  
344 TCATCGTCCGCGCAAGAAAGTTCCTCGCATTAATTTCAAGCTCCCGTGCA 403

239 TTACTAGTATTACCTGGACCTGGTCAATCCGACCGCGGCAATATAGCAAGTTCCTGG 298  
134 AATCTCTTCTATCAAGCGAAGATCCAAACTAAATGCTATGGCATACCAATCTTAC 193  
294 AATAGATCTGTAACAAAGTAAAGATCCCAATCTGAATACGCTGCTACTGATATTCG 358  
194 CTGATCTAATTCGACCTTAACTTATGTTAAGCTCCAAAGTTCGCAATTCGCA 253  
359 TCATCGCTCCCGGAGCAAGAAAGTCTCTCGCATTAATTTCAAAAGTTCGCA 418  
254 CCATTACACTAATCTGAGCAAGAAATTAATCTATACGCTGATGCTATTCGATCTTCA 313  
419 CTCTTCTCTGGGCTGAGCGGATTAATCTGATGCTATCTCTGATGCTATTCGATCT 478  
314 ATGCGAATAAGTCTGCT---TACCATATATTAATGATATTAACAAGCAAGCAAGTCT 370  
479 ATAGCAAGCTGAACCGGCTTACTTCTGCTAGGAGATTACGAGCGGCAATTCGCTG 538  
371 ATGTGGAGTAATCTCTTTGCTCAAGTTTCTAGTTTCTGCTGTAATGCTATTAATCA 430  
539 CTCTGTTCCGAGCGGCACTCTCAAGCAAAAGCACTGGAATATACGTAATTTAC 598  
431 ATAGCTTATATTCGACCATGAAAGAAAGCAAGTAATCAAGCAATCAAGTCTCAAT 490  
599 AGTCATCTGAGCAAGCAAGCGGATCATCCAGGCGGATCAATCCCGCAAGCAATCTG 658  
491 TGGGAATCAATCTAGCAGTACATGGAATAATCTCTGGAGTTGATCTATTCCTGCTG 550  
659 TGGGTATTGATCTGCTGAGCAGGAGCATGAGGCGGATCAATCCCGCAAGCAATCTG 715  
551 TAAAGCACTGAGCTTTTCTGCTAGCTAGCTCAATGCTGAGTCTGAGTCTGATCTG 610  
716 TTAAGAGCAAGCGGCTTCTGCTGCTGATCGGCTTCAAGTACGCGGCAAGCGGCTT 775  
611 TCAAGTACATAGCAAGCAAGTCAAGCAATTTTAAAGAGCAATCTTACCTGATCTCA 670  
776 TCGGCTACATTCAGAACTGCTCATCAAAAGCTTCCGCAAGTTCATTCGCAAGTA 835  
671 AAGTAATTAATTTGGAGGAGAAAGTGGGCAAAATCTCTGAGGCAATTCAGATCTCA 730  
836 AAGTCATCTGAGTCTGAGTTAATTTGAAAAAATTTCCACCCGCTTATGATGAGCA 895  
731 ATGGGCTTTTACCAA 746  
896 AGAAGCGTGTTCCTCAA 911

RESULT 15

US-09-792-793A-57  
; Sequence 57, Application US/09792793A  
; Patent No. US20020168370A1  
; GENERAL INFORMATION:  
; APPLICANT: McDonald, John R.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING SECONDARY TISSUE DAMAGE  
; FILE OF INVENTION: OTHER INFLAMMATORY CONDITIONS AND DISORDERS  
; FILE REFERENCE: 25020-6010  
; CURRENT APPLICATION NUMBER: US/09/792.793A  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 93  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 57  
; LENGTH: 999  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(999)  
; OTHER INFORMATION: Description of Artificial Sequence: Construct encoding chemokine  
; OTHER INFORMATION: toxin fusion protein MCP1-AM-SAPORIN  
; US-09-792-793A-54

Query Match 8.9% Score 84; DB 9; Length 999;

194 CTGATCTAATTCGACCCCTAAGTCTATTGCTTAAGCTCCAAAGTTCGCAAACTAA 253  
353 TCATCGCTCCCGGAGCAAGAAAGTCTCTGGCATTAATTTCAAAAGTTCGCTGGCA 412  
254 CCATTACACTAATCTGAGCAAGAAATTAATCTATACGCTGATGCTATTCGATCTTCA 313  
413 CTGTTCTCTGGGCTGAGCGGATTAATCTGATGCTATGTTGCTATCTGCGGATGNTA 472  
314 ATGCGAATAAGTCTGCT---TACCATATATTAATGATATTAACAAGCAAGCAAGTCT 370  
473 ATAGCAAGCTGAACCGGCTTACTTCTGCTAGGAGATTACGAGCGGCAATTCGCTG 532  
371 ATGTGGAGTAATCTTTGCTCAAGTTTCTAGTTTCTGCTGTAATGCTCTCAATTAATCA 430  
533 CTCTGTTCCGAGCGGCACTCTCAAGCAAAAGCACTGGAATATACGTAATTTAC 592  
431 ATAGCTTATATTCGACCATGAAAGAAAGCAAGTAATCAAGCAATCAAGTCTCAAT 490  
593 AGTCATCTGAGCAAGCAAGCGGATCATCCAGGCGGATCAATCCCGCAAGCAATCTG 652  
491 TGGGAATCAATCTAGCAGTACATGGAATAATCTCTGGAGTTGATCTATTCCTGCTG 550  
653 TGGGTATTGATCTGCTGAGCAGGAGCATGAGGCGGATCAACAAA---AAGCTCGCTGG 709  
551 TAAAGCACTGAGCTTTTCTGCTAGCTAGCTCAATGCTGAGTCTGAGTCTGATCTG 610  
710 TTAAGAGCAAGCGGCTTCTGCTGATGCTGCTGATCGGCTTCAAGTACGCGGCAAG 769  
611 TCAAGTACATAGCAAGCAAGTCAAGCAATTTTAAAGAGCAATCTTACCTGATCTCA 670  
770 TCGGCTACATTCAGAACTGCTCATCAAAAGCTTCCGCAAGTTCATTCGCAAGTA 829  
671 AAGTAATTAATTTGGAGGAGAAAGTGGGCAAAATCTCTGAGGCAATTCAGATCTCA 730  
830 AAGTCATCTGAGTCTGAGTTAATTTGAAAAAATTTCCACCCGCTTATGATGAGCA 899  
731 ATGGGCTTTTACCAA 746  
890 AGAAGCGTGTTCCTCAA 905

RESULT 14

US-09-792-793A-54  
; Sequence 54, Application US/09792793A  
; Patent No. US20020168370A1  
; GENERAL INFORMATION:  
; APPLICANT: McDonald, John R.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING SECONDARY TISSUE DAMAGE AND  
; FILE OF INVENTION: OTHER INFLAMMATORY CONDITIONS AND DISORDERS  
; FILE REFERENCE: 25020-6010  
; CURRENT APPLICATION NUMBER: US/09/792.793A  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 93  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 54  
; LENGTH: 999  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(999)  
; OTHER INFORMATION: Description of Artificial Sequence: Construct encoding chemokine  
; OTHER INFORMATION: toxin fusion protein MCP1-AM-SAPORIN  
; US-09-792-793A-54

Query Match 8.9% Score 84; DB 9; Length 999;

Best Local Similarity 47.3%; Pred. No. 1.3e-13;  
Matches 320; Conservative 0; Mismatches 350; Indels 6; Gaps 2;  
74 TAAATACGATCACCTTTGATGCTGGAATGCCACCAATTAACAATATGCACCTTTATCG 133



01-JAN-1988 (Rel. 06, Last sequence update)  
15-JUN-2002 (Rel. 41, Last annotation update)  
Agglutinin precursor (RCA) [contains: Agglutinin A chain (cRNA N-glycosidase) (EC 3.2.2.22); Agglutinin B chain].  
Ricinus communis (Castor bean).  
Lukarota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Malpighiales; Euphorbiaceae; Ricinus.  
NCBI\_TaxID=3988;  
[1]  
SEQUENCE FROM N.A.  
MEDLINE=86059449; PubMed=2999130;  
Roberts L.M., Lamb F.I., Pappin D.J.C., Lord J.M.:  
"The primary sequence of Ricinus communis agglutinin. Comparison with ricin."  
J. Biol. Chem. 260:15682-15686(1985).  
[2]  
SEQUENCE OF 303-564.  
TISSUE=Seed.  
C Araki T., Yoshioka Y., Funatsu G.:  
"The complete amino acid sequence of the B-chain of the Ricinus communis agglutinin isolated from large-grain castor bean seeds."  
Biochim. Biophys. Acta 872:277-285(1986).  
[3]  
SEQUENCE OF 303-337.  
MEDLINE=80178723; PubMed=6768555;  
Lin T.T.-S., Li S.S.-L.:  
"Purification and physicochemical properties of ricins and agglutinins from Ricinus communis."  
Eur. J. Biochem. 105:453-459(1980).  
[4]  
-!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one specific adenosine on the 28S rRNA.  
-!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.  
-!- SIMILARITY: CONTAINS 2 RICIN B-TYPE LECTIN DOMAINS.

\*\*\*\*\*  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
\*\*\*\*\*  
EMBL: M12089; AAA33869.1; -;  
EMBL: S40368; AAB22584.1; -;  
PIR: A24261; RLCSAG.  
HSP: P02879; IBK6.  
GlycoSuiteDB: P06750; -;  
InterPro: IPR001574; RIP.  
InterPro: IPR000772; Ricin\_B\_lectin.  
Pfam: PF00161; RIP; 1.  
Pfam: PF00652; Ricin\_B\_lectin; 6.  
PRINTS: P000396; SHIGARICIN.  
SMART: SM00458; RICIN; 2.  
PROSITE: PS00231; RICIN\_B\_LECTIN; 2.  
PROSITE: PS00275; SHIGA\_RICIN; 1.  
Plant defense: Hydrolase; Protein synthesis inhibitor; Toxin; Repeat: Glycoprotein; Lectin; Signal.  
SIGNAL  
1 24  
CHAIN 25 290  
25 290 AGGLUTININ A CHAIN.  
PROPEP 291 302  
302 LINKER PEPTIDE.  
CHAIN 303 564  
303 564 AGGLUTININ B CHAIN.  
DOMAIN 309 436  
436 RICIN B-TYPE LECTIN 1.  
DOMAIN 439 563  
563 RICIN B-TYPE LECTIN 2.  
REPEAT 319 361  
361 1-ALPHA.  
REPEAT 362 402  
402 1-BETA.  
REPEAT 405 437  
437 1-GAMMA.  
REPEAT 450 485  
485 2-ALPHA.  
REPEAT 489 528  
528 2-BETA.  
REPEAT 531 558  
558 2-GAMMA.  
ACT\_SITE 200 200  
200 BY SIMILARITY.  
DISULFID 282 306  
306 INTERCHAIN (BY SIMILARITY).

FT	DISULFID	322	341	BY SIMILARITY.
FT	DISULFID	365	382	BY SIMILARITY.
FT	DISULFID	453	466	BY SIMILARITY.
FT	DISULFID	492	509	BY SIMILARITY.
FT	CARBOHYD	34	34	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	259	259	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	397	397	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	437	437	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CONFLICT	331	331	F -> T (IN REF. 2).
FT	CONFLICT	362	362	N -> D (IN REF. 2).
FT	CONFLICT	374	374	R -> G (IN REF. 2).
FT	CONFLICT	404	404	R -> T (IN REF. 2).
FT	CONFLICT	552	552	F -> V (IN REF. 2).
SQ	SEQUENCE	564 AA:	62851 MW:	D455F2A72F609759 CRC64:

Query Match 16.8%; Score 271.5; DB 1; Length 564;  
Best Local Similarity 28.7%; Pred. No. 4e-14;  
Matches 77; Conservative 47; Mismatches 107; Indels 37; Gaps 7;

QY	28	ITFDAGNATINKYATFMSURNOAKDKPKLCYGIPIMLPDTNSTP---KYLLVKILOGANIK	84
DB	33	INFTIADATVESVTNFTRAVRSILTTGADVRRHEIPVLPNRVGLPISQRFILVELSNIHEL	92
QY	85	TITLMLRRNNLYVNG-----YSDPENGKCRHYHFNIDITSTERTTOVENTLSSSS	134
DB	93	SVTLALDYTNAYVYVCGRAGNSAYFFHPDNOEDADALPHLP-----TDQVN-----	137
QY	135	SRVAMSYNNSLYPTMEKKAENSRNOVQIGIQLSSDGIKI--SGVDSFPVKTEAPFLI	192
DB	138	---SETAFGCGNYDRLEQJGL--RENIEIGTGTPLEDAISALYYSTCGTQIPILARSFM	192
QY	193	VAIQMSVDAARFKYIENOVKT--NFNRAFYPPDPKVINLEEKWKGLKSEAIHNAKNGALPKP	250
DB	193	VCIONISEAAHQVIEGEMRTIRYNRSAPDPSVITLNSWGRSLTAIOESNOGAFASP	252
QY	251	LELVDAKGTWIKVLKRVDEINRDVALIKY	278
DB	253	IQLORNGSKFNVDVDSILIPITIALMVY	280

Search completed: May 28, 2003, 09:56:38

Job time : 27 secs





67 LTTGADVREHPVLPNVRGLPINORFIFVLSNHAELSVTLALDVTNAYVYCYR---AGN 123  
108 KCRY-----HIFNDITSTERTDVENTLCSSSSSRVAMSYNLSLYPTMEKA 154  
124 SAYFFHPDQDEAETHLF-----TDVQN-----RYTAFGCGN--YDRLEQLA 165  
155 EVNSRNVOLGILQLSSDGKI-----SGVDSFPVKTEAFELLVAIQMSEARFKYIENQ 210  
166 G-NURENTELGNGLEEAISALYYSTGTLPTLARSF--IICIMISEARFPYIEGE 222  
211 VKT--NENRATPPKPVNLEKKGKISEAHNAKNGALPKPLELVDAKGKWIIVLRVDE 268  
223 WRTRRYNRGAPDSVITLNSWGRSLSTQESNGAFASPIQLORRNGSKFSYDVYSI 282  
269 INROVALKY 278  
283 LIPITIALMVY 292

## RESULT 12

RIP1\_TRIAN STANDARD: PRT: 294 AA.  
P56626; Q9ZQY7;  
15-DEC-1998 (Rel. 37, Created)  
30-MAY-2000 (Rel. 39, Last sequence update)  
15-JUN-2002 (Rel. 41, Last annotation update)  
Type I ribosome-inactivating protein trichoanguina precursor (rRNA N-glycosidase) (EC 3.2.2.22) (RIP) (Trichoanguin).  
TCA.  
Trichosanthes anguina (Snake gourd).  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.  
NCBI\_TaxID=50544;  
[1]  
SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND CHARACTERIZATION.  
STRAIN=CV. Anguina; TISSUE=Seed.  
MEDLINE=99132006; PubMed=993118;  
Chow L.-P., Chou M.-H., Ho C.-Y., Chuang C.-C., Pan F.-M., Wu S.-H., Lin J.-Y.;  
"Purification, characterization and molecular cloning of trichoanguin, a novel type I ribosome-inactivating protein from the seeds of Trichosanthes anguina.";  
Biochem. J. 338:211-219(1999).  
[2]  
SEQUENCE OF 20-264.  
TISSUE=Seed;  
Chow L.-P., Kamo M., Lin J.-Y., Wang S.-H., Ueno Y., Tsugita A.;  
"Amino acid sequence of trichoanguina, a ribosomal-inactivating protein from Trichosanthes anguina seeds.";  
J. Biomed. Sci. 3:178-186(1996).  
[3]  
-!- FUNCTION: INHIBITS PROTEIN SYNTHESIS BY DEPURINATING 28S RRNA IN RIBOSOMES.  
-!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one specific adenosine on the 28S rRNA.  
-!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
TYPE I RIP SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

EMBL: AF055086; AAD02686.1;  
HSSP: P33185; LBRY.  
InterPro: IPR001574; RIP.  
Pfam: PF00161; RIP; 1.  
PRINTS: PR003396; SHIGARICIN.  
PROSITE: PS003275; SHIGA-RICIN; FALSE\_NEG.  
Plant defense; Protein synthesis inhibitor; Hydrolase; Toxin;

KW Glycoprotein: Signal. 19  
FT CHAIN 1 20 264  
FT TYPE I RIBOSOME-INACTIVATING PROTEIN  
FT TRICHOANGUINA.  
FT  
FT PROPEP 265 294  
FT ACT\_SITE 177 177  
FT ACT\_SITE 180 180  
FT CARBOHYD 70 70  
FT CARBOHYD 220 220  
FT CONFLICT 51 51  
FT CONFLICT 65 65  
FT CONFLICT 84 84  
FT CONFLICT 152 152  
FT CONFLICT 174 174  
FT CONFLICT 245 245  
SQ SEQUENCE 294 AA: 32234 MW: 32234 MW: DA4F8B7CE3290994 CRC64;  
Query Match 18.6%; Score 300; DB 1; Length 294;  
Best Local Similarity 30.0%; Pred. No. 9.3e-17;  
Matches 85; Conservative 57; Mismatches 113; Indels 28; Gaps 10;  
QY 12 IAMLIAAPTSTCAINTITFDAGNATINKYATFMESLRNQAKDPKLCYGIPLPDTNSTP 71  
DB 8 LAISLGSPT---AIGDSFSLSTATKKSYSSEITQLDRALPTQCTVC-GIPLEPSTANGS 63  
QY 72 KYL-LVKLOGANLKTITLMLRRNLLYVNGY--SDPENGKCRVHIFNDITSTERTOVNTL 129  
DB 64 QMFRFNTNTNDETIVAVVNTVIVAYRADAVS-----YFED---TPAENAKLIF 134  
QY 130 CSSSSSRVAMSYNLSYPTMEKKAENSRNQVQIGIQISDIDGKISGVNSFPVKTKAF 189  
DB 115 ACTKTVKLPYSUNYQKLSVVGKQ-----RDMIELGIPALSS---AITNMVYVDYSTAA 166  
QY 190 FILVALQMVSEARFKYIENQVKNFENAFYDDPKVINLEKWKGLSEAIHNAK---NCA 246  
DB 167 ALLVLIOCTAAARYKYIEQVSSH[SSNFYPNQAVISLENKNGALSKQIQIOTAMRTGIGQ 226  
QY 247 LPKPLFLVDAGCKTKWIVLRVDE--INROVALKYVNGTCQTTY 287  
DB 227 FENPVELYNPQITRESVTNTSAGVVGKNGIKULLLYKASVSGSEY 269

## RESULT 13

RIP3\_SAPOF  
ID RIP3\_SAPOF STANDARD: PRT: 236 AA.  
AC P27560;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Ribosome-inactivating protein saporin-3 (SAP-3) (SO-3) (rRNA N-glycosidase) (EC 3.2.2.22) (Fragment).  
GN SAP3.  
OS Saponaria officinalis (Common soapwort).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Caryophyllales; Caryophyllaceae; Saponaria.  
OX NCBI\_TaxID=3572;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92049247; PubMed=1719367;  
RA Fordham-Skelton A.P., Taylor P.E., Hartley M.R., Croy R.K.D.;  
RT "Characterisation of saporin genes: in vitro expression and ribosome inactivation";  
RL Mol. Gen. Genet. 229:460-466(1991).  
CC -!- FUNCTION: RIBOSOME-INACTIVATING PROTEIN OF TYPE 1. INHIBITS PROTEIN SYNTHESIS IN ANIMAL CELLS. USEFUL AS IMMUNOTOXIN FOR PHARMACOLOGICAL APPLICATIONS.  
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one specific adenosine on the 28S rRNA.  
CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
CC TYPE I RIP SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration

MEDLINE=91352004; PubMed=1881881;

Katzin B.J., Collins E.J., Robertus J.D.;

"Structure of ricin A-chain at 2.5 A.";

Proteins 10:251-259(1991).

[10]

X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF B-CHAIN.

MEDLINE=91352005; PubMed=1881882;

Rutember E., Robertus J.D.;

"Structure of ricin B-chain at 2.5-A resolution.";

Proteins 10:260-269(1991).

[11]

X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF A-CHAIN.

MEDLINE=95082010; PubMed=7990130;

Weston S.A., Tucker A.D., Thatcher D.R., Derbyshire D.J.,

"X-ray structure of recombinant ricin A-chain at 1.8-A resolution.";

J. Mol. Biol. 244:410-422(1994).

[12]

X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF MUTANT HIS-215.

MEDLINE=96374222; PubMed=8780513;

Day P.J., Ernst S.R., Frankel A.E., Monzingo A.F., Pascal J.M.,

"Structure and activity of an active site substitution of ricin A

chain.";

Biochemistry 35:11098-11103(1996).

[13]

X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF A-CHAIN.

MEDLINE=97210820; PubMed=9086280;

Yan X., Hollis T., Svinth M., Day P., Monzingo A.F., Milne G.W.,

"Structure-based identification of a ricin inhibitor.";

J. Mol. Biol. 266:1043-1049(1997).

[14]

MUTAGENESIS.

MEDLINE=93165632; PubMed=1287657;

Kin Y., Robertus J.D.;

"Analysis of several key active site residues of ricin A chain by

mutagenesis and X-ray crystallography.";

Protein Eng. 5:775-779(1992).

[15]

1- FUNCTION: Ricin is highly toxic to animal cells and to a less extent to plant cells. The A chain is responsible for inhibiting protein synthesis through the catalytic inactivation of 60S ribosomal subunits. It acts as a glycosylase that removes a specific adenine residue from an exposed loop of 28S ribosomal RNA. As this loop is involved in the binding of elongation factors, the modified ribosomes are unable to support protein synthesis. The A chain can inactivate a few thousand ribosomes per minute, thus inactivating them faster than the cell can make new ones. A single A-chain molecule can therefore kill an animal cell. The B chain binds to cell receptors and facilitates the entry into the cell of the A chain; B chains are also responsible for cell agglutination (lectin activity). It binds to beta-D-galactopyranoside moieties.

1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one specific adenosine on the 28S rRNA.

1- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.

1- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS (ALPHA, BETA, GAMMA).

1- PTM: THE MAJOR A-CHAIN IS GLYCOSYLATED ONLY IN POSITION 45 AND THE

MINOR A-CHAIN IS GLYCOSYLATED ALSO IN POSITION 271.

1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RIBOSOME-

INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.

1- SIMILARITY: CONTAINS 2 RICIN B-TYPE LECTIN DOMAINS.

1- CAUTION: REF.4 AND REF.5 SEQUENCES HAVE A HIGH NUMBER OF CONFLICTS

WITH THE SEQUENCE TRANSLATED FROM DNA (REF.1, REF.2 AND REF.3).

[16]

THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```

CC      EMBL: X03179; CAA26939.1; -
DR      EMBL: X52908; CAA37095.1; -
DR      EMBL: X02388; CAA26230.1; -
DR      EMBL: A12892; CAAU1058.1; -
DR      PIR: A24041; RLCSD.
DR      PDB: 2AAL; 31-JAN-94.
DR      PDB: 1APG; 31-JAN-94.
DR      PDB: 1FMP; 31-OCT-93.
DR      PDB: 1IFS; 14-JAN-98.
DR      PDB: 1IFU; 14-JAN-98.
DR      PDB: 1LTC; 31-OCT-93.
DR      PDB: 1LBS; 16-JUN-97.
DR      PDB: 1OBT; 16-JUN-97.
DR      PDB: 1BRS; 02-SEP-98.
DR      PDB: 1BR6; 02-SEP-98.
DR      GlycoStatedR: P02879; -
DR      InterPro: IPR001574; RIP.
DR      InterPro: IPR000772; Ricin_B_lectin.
DR      Pfam: PF00161; RIP; 1.
DR      Pfam: PF00552; Ricin_B_lectin; 6.
DR      PRINTS: PM00396; SHIGARICIN.
DR      SMART: SM00458; RICIN; 2.
DR      PROSITE: PS00275; SWIGA_RICIN; 2.
DR      PROSITE: PS02311; RICIN_B_LECTIN; 2.
DR      Plant defense: Hydrolase; Protein synthesis inhibitor; Toxin; Repeat.
DR      Glycoprotein: Lectin; Signal; 3D-structure.
DR      SIGNAL 1 35
FT      CHAIN 36 302      RICIN A CHAIN.
FT      PEPTIDE 303 314      LINKER PEPTIDE.
FT      CHAIN 315 576      RICIN B CHAIN.
FT      DOMAIN 321 448      RICIN B-TYPE LECTIN 1.
FT      DOMAIN 451 575      RICIN B-TYPE LECTIN 2.
FT      REPEAT 331 373      1-ALPHA.
FT      REPEAT 374 414      1-BETA.
FT      REPEAT 417 449      1-GAMMA.
FT      REPEAT 462 497      2-ALPHA.
FT      REPEAT 501 540      2-BETA.
FT      REPEAT 543 570      2-GAMMA.
FT      ACT_SITE 212 212      INTERCHAIN.
FT      DISULFID 294 318
FT      DISULFID 334 353
FT      DISULFID 377 394
FT      DISULFID 465 478
FT      DISULFID 504 521
FT      CARBOHYD 45 45      N-LINKED (GLCNAC. . .).
FT      CARBOHYD 271 271      /FTID-CAR_000080.
FT      FT      N-LINKED (GLCNAC. . .); IN MINOR A-CHAIN
FT      FT      VARIANT.
FT      FT      /FTID-CAR_000081.
FT      CARBOHYD 409 409      N-LINKED (GLCNAC. . .).
FT      CARBOHYD 449 449      N-LINKED (GLCNAC. . .).
FT      CONFLICT 76 76      E -> D (IN REF. 3).
FT      CONFLICT 551 551      A -> R (IN REF. 3).
FT      STRAND 43 47
FT      TURN 49 50
FT      HELIX 53 67
FT      STRAND 73 74
FT      TURN 75 76
FT      STRAND 77 79
FT      TURN 88 90

```

Query Match 18.98; Score 305.5; DB 1; Length 576;

Best Local Similarity 28.74; Pred. No. 8.9e-17;

Matches 89; Conservative 48; Mismatches 108; Indels 65; Gaps 11;

OY 10 TLIAWLAIAAPTSTCAIMT-----ITFDAGNATINKVATFMSLRNQ 50

| : : | : : | : : | : : | : : | : : | : : | : : | : : |

Db 7 TIVIMYAVATLFCGTSQSWFTLENNIIFE-PIINFTTAGATVQSYNFIKAVGR 66

| : : | : : | : : | : : | : : | : : | : : | : : | : : |

OY 51 AKDPKIKCYGIPHLIPOTNSTP---KYLIIVKIQGANLKTITLMLRRNLYVMGYSNIPFN 107

: : : | : : | : : | : : | : : | : : | : : | : : | : : |

Gene 134:223-227(1993).  
[2]  
SEQUENCE OF 47-93.  
TISSUE=Seed;  
MEDLINE=89326691; PubMed=2753596;  
Montecucchi P.C., Lazzarini A.M., Barbieri L., Stirpe F., Soria M.,  
Lapli D.;  
Int. J. Pept. Protein Res. 33:263-267(1989).  
[3]  
X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).  
MEDLINE=9533189; PubMed=7608981;  
Iosur M.V., Nair B., Sathyanarayana P., Misquith S., Surolla A.,  
Kannan K.K.;  
"X-ray structure of gelonin at 1.8-A resolution.";  
J. Mol. Biol. 250:368-380(1995).  
[4]  
- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
specific adenosine on the 28S rRNA.  
[5]  
- SUBUNIT: HOMODIMER.  
[6]  
- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
[7]  
TYPE 1 RIP SUBFAMILY.  
[8]  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL Outstation  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
[9]  
EMBL: U12243; AAA16312.1; -  
PIR: S16489; S16489.  
[10]  
ISSP: P09989; LMRJ.  
[11]  
Inter-Pro: IP0001574; RIP.  
[12]  
Pfam: PF00161; RIP; 1.  
[13]  
PRINTS: P000396; SHIGARICIN.  
[14]  
PROSITE: PS00275; SHIGA-RICIN; FALSE\_NEG.  
[15]  
Plant defense; Protein synthesis inhibitor; Hydrolase; Toxin; Signal;  
Glycoprotein.  
[16]  
SIGNAL: 1 26 POTENTIAL.  
[17]  
PROPEP 27 46  
CHAIN 47 297 RIBOSOME-INACTIVATING PROTEIN GELONIN.  
[18]  
PROPEP 298 316 MISSING IN MATURE PROTEIN.  
[19]  
DISULFID 90 96  
CARBOHYD 235 235 N-LINKED (GLCNAC. . .).  
[20]  
ACT\_SITE 212 212  
[21]  
CONFLICT 90 90 C -> K (IN REF. 2).  
[22]  
CONFLICT 93 93 P -> D (IN REF. 2).  
[23]  
SEQUENCE 316 AA: 35418 MW: 125273710901885 CRC64:  
[24]  
Query Match: 20.4%; Score 330; DB 1; Length 316;  
Best Local Similarity 30.8%; Pred. No. 4.2e-19;  
Matches 97; Conservative 60; Mismatches 122; Indels 36; Gaps 11;  
[25]  
7 VVTLIAWLAAPTS-----TCAINTTFDAGNATINKYATWESLRNQAOKPLKCY 59  
[26]  
23 IVLGSTARIFSLPTNDEETSKTGLGLOTVSEFTKGATVITVYVFNELNRLVKL-PEGNSH 81  
[27]  
60 GIPMLPDTNSTPK--YLLVKKLQGANLKTITLMLRNRLNLYVMGSDPPFNGKRCRYHIENDI 117  
[28]  
82 GIPLLRKKCDPPGKCFVLVALSNONGLAETADVTSVYVVGQ-----VNRNSYFFKDA 136  
[29]  
118 TSTERTDV-ENTLCSSSSSRVAMSTYNLSNLYPTMKKAEVNSRNOVQLGIOTLSSDICKI 176  
[30]  
137 PDAAVEGLFKNT-----IKTRLFGSGYPSLE--GEKAYRETTDLGIEPLRIGIKKL 186  
[31]  
177 --SGVDSPPVTEAFLLVATOMYSEARFYKYNQVKTNNRFRYPDPKVINLEKWK 234  
[32]  
187 DENAIDNYKPTIEASSLLVIVOMYSEARFTFIENQIRNFOORIRPANNTISLENKWK 246  
[33]  
235 ISEATH-NAKNAKPKLLELDVAKGKTIVILRVDSINRVDALLKYNVNGTCOTTYGNAMFS 293  
[34]  
247 LSFQRTSGANMFSEAVELELRANGKKYVYTAQVQPKIALLEKVDKPKT----SLAA 302

OY 294 QVIITTYNYMSNLG 308  
DB 303 ELIIO---NYESLVG 314  
RESULT 11  
RICI, RICO  
ID RICI, RICO STANDARD: PRT: 576 AA.  
AC P02679; P02880;  
DT 21-JUL-1985 (Rel. 01, Created)  
DT 13-AUG-1987 (Rel. 05, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Ricin precursor [Contains: Ricin A chain (rRNA N-glycosidase)  
(EC 3.2.2.22); Ricin B chain].  
DE Ricinus communis (Castor bean).  
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicot; Rosidae;  
OX eucosids 1; Malpighiales; Euphorbiaceae; Ricinus.  
OX NCBI\_TaxID=3988;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86067214; PubMed=2999712;  
RA Halling K.C., Halling A.C., Murray E.E., Ladin B.F., Houston L.L.,  
Weaver R.F.;  
RT "Genomic cloning and characterization of a ricin gene from Ricinus  
communis";  
RL Nucleic Acids Res. 13:8019-8033(1985).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92163016; PubMed=1371405;  
RA Tregear J.W., Roberts L.M.;  
RT "The lectin gene family of Ricinus communis: cloning of a functional  
ricin gene and three lectin pseudogenes";  
RL Plant Mol. Biol. 18:515-525(1992).  
RN [3]  
RP SEQUENCE OF 12-576 FROM N.A.  
RX MEDLINE=85179479; PubMed=3838723;  
RA Lamb A., Roberts L.M., Lord J.M.;  
RT "Nucleotide sequence of cloned cDNA coding for preproricin";  
RL Eur. J. Biochem. 148:265-270(1985).  
RN [4]  
RP SEQUENCE OF 36-302  
RA Toshioka S., Funatsu G., Funatsu M.;  
RT "Isolation and sequences of peptic peptides, and the complete  
sequence of 11e chain of ricin-D";  
RL Agric. Biol. Chem. 42:1267-1274(1978).  
RN [5]  
RP SEQUENCE OF 315-576.  
RA Funatsu G., Kimura M., Funatsu M.;  
RT "Primary structure of Ala chain of ricin D";  
RL Agric. Biol. Chem. 43:2221-2224(1979).  
RN [6]  
RP CARBOHYDRATE-LINKAGE SITES, AND PARTIAL SEQUENCE.  
RX MEDLINE=90344223; PubMed=1368517;  
RA Kimura Y., Kusoku H., Tada M., Takagi S., Funatsu G.;  
RT "Structural analyses of sugar chains from ricin A-chain variant";  
RL Agric. Biol. Chem. 54:157-162(1990).  
RN [7]  
RP REVIEW.  
RX MEDLINE=21480122; PubMed=11595634;  
RA Olsnes S., Kozlov J.V.;  
RL Toxicon 39:1723-1728(2001).  
RN [8]  
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).  
RX MEDLINE=87165983; PubMed=3558397;  
RA Monfort W., Villafraanca J.E., Monzingo A.F., Ernst S.R., Katzin B.,  
Rutember E., Xuong N.H., Hamlin R., Robertus J.D.;  
RT "The three-dimensional structure of ricin at 2.8 A";  
RL J. Biol. Chem. 262:5398-5403(1987).  
RN [9]  
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF A-CHAIN.

RT RT FUNCTION: RIBOSOME-INACTIVATING PROTEIN OF TYPE 1, INHIBITS  
 CC CC PROTEIN SYNTHESIS IN ANIMAL CELLS (BY SIMILARITY).  
 CC CC CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
 CC CC specific adenosine on the 28S rRNA.  
 CC CC SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 CC CC TYPE 1 RIP SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: X69131; CAA48885.1;  
 CC EMBL: X69132; CAA48886.1;  
 CC HSSP: P10297; IQCG.  
 CC InterPro: IPR001574; RIP.  
 CC Pfam: PF00161; RIP.  
 CC PROSITE: PS00275; SHIGA\_RICIN; 1.  
 CC Plant defense: Protein synthesis inhibitor; Hydrolase; Toxin;  
 CC Multigene family.  
 CC ACT\_SITE 176 176 BY SIMILARITY.  
 CC SEQUENCE 253 AA: 28556 MW: 539693E13D0F594D CRC64:

Query Match 20.8% Score 337; DB 1; Length 253;  
 Best Local Similarity 35.5% Pred. No. 8.8e-20;  
 Matches 93; Conservative 50; Mismatches 99; Indels 20; Gaps 9;

25 INTITDAGNATINKYATPMESLRNOKAPKLCYCIPM-LPDNSTPKYLLVLOGANL 83  
 1 VTSITLDLVNPTAGOVSSFDVKIRNKPDLKYGCTDIAGVPPSKKFLRINFQSSR- 59  
 84 KTTILMLRNNLYVMGYSDPFNGKNC-R-YHFNIDTSTERTDVENTLCSSSSRVAMSN 142  
 60 GTVSLGCKNDLVVAYLAMDNNTNVRAYFRSEITSAELT-----ALFPEATANOKALE 115  
 143 YNSLYPTMEKKAEEV-----NSRNQVQLGQILSSDICKISGVDSFPVKTEAFELLVAIQMV 198  
 116 YTEDQSIENKNAQITGGDKSRKELGIGDILLTSMNAV-N-KKRVYKNEARELLIAQMT 174  
 199 SEAAEFKYTENOVKNFENRAEYDPDKVINLEEKWCKISEAIIH-NAKNGALPKPLGLVDIAK 257  
 175 AEVAFRYITONLVTKNFPKNSKVKVIOFEVSWRKISTAIYGDANKGVFNKDYDFGCK 234  
 258 GTKWIVLRVDEINROVALLKYV 279  
 235 -----VRQVKDL--QNGLLMYL 249

RESULT 9  
 ID RIP7 SAPOF STANDARD: PRT: 253 AA.  
 AC P31186;  
 DT 6-OCT-2001 (Rel. 40, Created)  
 DT 6-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Ribosome-inactivating protein saporin-7 (SO-7) (SAP-7) (rRNA N-  
 DE glycosidase) (EC 3.2.2.22).  
 DE (SAP7).

GN Saponaria officinalis (Common soapwort).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Caryophyllales; Caryophyllaceae; Saponaria.  
 CC NCBI\_TaxID=3572;  
 CC [1]  
 CC SEQUENCE FROM N.A.  
 CC TISSUE=Leaf;  
 CC MEDLINE=93203250; PubMed=8454624;  
 CC Barthelemy I., Martineau D., Ong M., Matsunami R., Ling N.,  
 CC Benatti L., Cavallaro U., Soris M., Lappi D.A.;  
 CC The expression of saporin, a ribosome-inactivating protein from the

RT plant Saponaria officinalis, in *Escherichia coli*.  
 RL J. Biol. Chem. 268:6541-6548(1993).  
 CC -!- FUNCTION: RIBOSOME-INACTIVATING PROTEIN OF TYPE 1, INHIBITS  
 CC CC PROTEIN SYNTHESIS IN ANIMAL CELLS (BY SIMILARITY).  
 CC CC CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
 CC CC specific adenosine on the 28S rRNA.  
 CC CC SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 CC CC TYPE 1 RIP SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: X69131; CAA48888.1;  
 CC HSSP: P10297; IQCG.  
 CC InterPro: IPR001574; RIP.  
 CC Pfam: PF00161; RIP.  
 CC PROSITE: PS00275; SHIGA\_RICIN; 1.  
 CC Plant defense: Protein synthesis inhibitor; Hydrolase; Toxin;  
 CC Multigene family.  
 CC ACT\_SITE 176 176 BY SIMILARITY.  
 CC SEQUENCE 253 AA: 28553 MW: 161319102AB20539 CRC64:

Query Match 20.8% Score 336; DB 1; Length 253;  
 Best Local Similarity 35.5% Pred. No. 1.1e-19;  
 Matches 93; Conservative 51; Mismatches 98; Indels 20; Gaps 9;

QY 25 INTITDAGNATINKYATPMESLRNOKAPKLCYCIPM-LPDNSTPKYLLVLOGANL 83  
 DB 1 VTSITLDLVNPTAGOVSSFDVKIRNKPDLKYGCTDIAGVPPSKKFLRINFQSSR- 59  
 QY 84 KTTILMLRNNLYVMGYSDPFNGKNC-R-YHFNIDTSTERTDVENTLCSSSSRVAMSN 142  
 DB 60 GTVSLGCKNDLVVAYLAMDNNTNVRAYFRSEITSAELT-----ALFPEATANOKALE 115  
 QY 143 YNSLYPTMEKKAEEV-----NSRNQVQLGQILSSDICKISGVDSFPVKTEAFELLVAIQMV 198  
 DB 116 YTEDQSIENKNAQITGGDKSRKELGIGDILLTSMNAV-N-KKRVYKNEARELLIAQMT 174  
 QY 199 SEAAEFKYTENOVKNFENRAEYDPDKVINLEEKWCKISEAIIH-NAKNGALPKPLGLVDIAK 257  
 DB 175 AEVAFRYITONLVTKNFPKNSKVKVIOFEVSWRKISTAIYGDANKGVFNKDYDFGCK 234  
 QY 258 GTKWIVLRVDEINROVALLKYV 279  
 DB 235 -----VRQVKDL--QNGLLMYL 249

RESULT 10  
 ID RIP7 SAPOF STANDARD: PRT: 316 AA.  
 AC P31186;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Ribosome-inactivating protein saporin precursor (rRNA N-glycosidase)  
 DE (EC 3.2.2.22).  
 DE GEL.

GN Gelonium multiflorum (Euphorbiaceae himalayana).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eucosids I; Malpighiales; Euphorbiaceae; Gelonium.  
 CC NCBI\_TaxID=39795;  
 CC [1]  
 CC SEQUENCE FROM N.A.  
 CC MEDLINE=94085781; PubMed=7916721;  
 CC Nolan P.A., Garrison D.A., Retter M.;  
 CC "Cloning and expression of a gene encoding gelonin, a ribosome-  
 CC inactivating protein from *Gelonium multiflorum*."

Multigene family.  
 SIGNAL 1 24  
 CHAIN 25 292  
 ACT\_SITE 200 200  
 VARIANT 72 72  
 SEQUENCE 292 AA: 32810 MM: FAL43C0E1B88976 CRC64:

Query Match 23.0%: Score 372: DB 1: Length 292:  
 Best Local Similarity 35.4%: Pred. No. 1.8e-22:  
 Matches 108: Conservative 52: Mismatches 107: Indels 38: Gaps 12:

4 MLVVVVTLIAWLI---AAPTSTCAINTITDAGNATINKYATFMSLRNOAKDKPKLCYGG 60  
 1 MKIYVATIAWILLOFSAWTTDAVTSITLVLNPTAGQYSSFDKIRNNVKDPNLYKGG 60  
 61 IPM-LPDTNSTPKYLLVKLGQANIKTITMLRRNNLYVMGSDPENGKCR-YHFNDDI 118  
 61 TDIAVIGPPSKDKELRINFQSSR-GTVSLGLKRONLYVAVAYLANDNTNVRAYFKSEIT 119  
 119 STERTDVENTLCSSSSSRVAMSYNSLYPTMEKAEV-----NSRNOVOLGIGIQL---SS 171  
 120 SAELUT-----ALFPPEATTANOKALEYTDYOSIERNAQITQCKSRKELGLGIDLLTME 175  
 172 DIGKISGVDPFPVKTAEFFLLVAIOMVSEARFYIENQVKTNFNRAFYDPKVINLEEK 231  
 176 AVNKKARV---VKNAREFLIAIOMTAEVARYIQLNLTKNFKNFSDNKNVYQFEVS 231  
 232 WCKTSEAIH-NAKNGALPKPLELVDAKGTWKIVLRVDEINRDVALLKYV-----NG 281  
 232 WRKISTAIYGAKNKGVNKNYDFGFK-----VRQVKOL--QMGLLMLYGLRKPSSNEANS 284  
 282 TCOTT 286  
 285 TAYAT 289

DIACA STANDARD: PRT: 293 AA.  
 P24476:  
 01-MAR-1992 (Rel. 21, Created)  
 01-DEC-1992 (Rel. 24, Last sequence update)  
 15-JUN-2002 (Rel. 41, Last annotation update)  
 Antiviral protein DAP-30 precursor (Ribosome-inactivating protein)  
 (RNA N-glycosidase) (EC 3.2.2.22) (Dianthin 30).  
 Dianthus caryophyllus (Carnation) (Clove pink).  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Caryophyllales; Caryophyllaceae; Dianthus.  
 NCBI\_TaxID=3570;  
 [1]  
 SEQUENCE FROM N.A.  
 MEDLINE=91355219; PubMed=1840496;  
 Legname G., Bellio P., Gromo G., Modena D., Keen J.N., Roberts L.M.,  
 Lord J.M.:  
 "Nucleotide sequence of cDNA coding for dianthin 30, a ribosome  
 inactivating protein from Dianthus caryophyllus.";  
 Biochim. Biophys. Acta 1090:119-122(1991).  
 [2]  
 SEQUENCE OF 24-82.  
 TISSUE=Leaf;  
 MEDLINE=92037998; PubMed=1936243;  
 Lee-Huang S., Kung H.-F., Huang P.L., Huang P.L., Li B.-Q., Huang P.,  
 Huang H.-I., Chen H.-C.:  
 "A new class of anti-HIV agents: CAP31, DAPS 30 and 32.";  
 FEBS Lett. 291:139-144(1991).  
 HIGH ANTIVIRAL POTENCY AND LOW TOXICITY TO NORMAL CELLS IN CULTURE  
 AND TO INTACT ANIMALS. CAPABLE OF INHIBITING HIV-1 INFECTION AND  
 REPLICATION.  
 CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
 specific adenosine on the 28S rRNA.  
 SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.

CC TYPE 1 RIP SUBFAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See [http://www.isb-sib.ch/announcements/](http://www.isb-sib.ch/announcements)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC

CC EMBL: X59260; CAA41953.1; ALT\_SRPQ.  
 CC PIR: S17519; S17519.  
 CC HSP: O03464; IAPA.  
 CC InterPro: IPR001574; RIP.  
 CC Pfam: PF00161; RTP.1.  
 CC PRINTS: PRO0396; SHIGARICIN.  
 CC PROSITE: PS00275; SHIGA\_MICIN.1.  
 CC Plant defense: Antiviral; Protein synthesis inhibitor; Hydrolase;  
 CC Toxin; Signal.  
 CC

FT SIGNAL 1 23  
 FT CHAIN 24 293 ANTIVIRAL PROTEIN DAP-30.  
 FT ACT\_SITE 200 200 BY SIMILARITY.  
 SO SEQUENCE 293 AA: 32717 MM: AC85AAAC8725D4F2 CRC64:

Query Match 21.4%: Score 345.5; DB 1: Length 293:  
 Best Local Similarity 34.6%: Pred. No. 2.3e-20:  
 Matches 100: Conservative 52: Mismatches 108: Indels 29: Gaps 12:

QY 4 MLVVVVTLIAWLIAPTS--TCAITNTITFDIAGNATINKYATFMSLRNOAKDKPKLCYGG- 60  
 DB 1 MKIYVATIAWILLOFSAWTTDAATAYTLNLANPSSASQYSSFLDQIRNNVRDTSL-1YGG 59  
 QY 61 --IPMLPDTNSTPKYLLVKLGQANIKTITMLRRNNLYVMGSDPENGKCR-YHFNDDI 117  
 DB 60 TDVAVIGCAPSTIDKELNFOGPK-GTVSLGIRRENTYVAVAYLANDNANVRAYFKNQI 118  
 QY 118 TSTERTDVENTLCSSSSSRVAMSYNSLYPTMEKAEV-----NSRNOVOLGIGIQLSSDI 173  
 DB 119 TSAELTALFPEVVVANOKQ----LEYGEDYQAIENKAKITTDGDSRKELGLGIGINL--- 171  
 QY 174 KISGVIS--FPVKTAEFFLLVAIOMVSEARFYIENQVKTNFNRAFYDPKVINLEEK 231  
 DB 172 TMIDGVNKKVVRVDEAREFLIAIOMTAEVARYIQLNLTKNFKNFSDNKNVYQFEVS 231  
 QY 232 WCKTSEATH-NAKNGALPKPLELVDAKGTWKIVLRVDEINRDVALLKYV 279  
 DB 232 WSKISTALFGCKNGVKNKYDFGFK-----VRQAKDL--QMGLLKYL 273

## RESULT 8

RIP5\_SAPOF STANDARD: PRT: 253 AA.  
 AC Q41389:  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Ribosome-inactivating protein saporin-5 (EC 3.2.2.22) (SAP-5) (rRNA N-  
 DE glycosidase).  
 GN SAP5.  
 OS Saponaria officinalis (Common soapwort).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Caryophyllales; Caryophyllaceae; Saponaria.  
 NCBI\_TaxID=3572;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=Leaf;  
 RX MEDLINE=93203250; PubMed=8454624;  
 RA Barthelemy T., Martineau D., Ong M., Matsunami K., Ling N.,  
 Benatti L., Cavallaro U., Sorla M., Lappi D.A.:  
 "The expression of saporin, a ribosome-inactivating protein from the  
 plant Saponaria officinalis, in Escherichia coli.";  
 J. Biol. Chem. 268:6541-6548(1993).





Matches 217; Conservative 14; Mismatches 30; Indels 1; Gaps 1;  
25 INTITFDAGNATINATYATMESLRNOAKDPKLCYCIPLMPTNTPKYLVLKQGANLK 84  
|||||  
1 INTITFDAGNATINATYATMESLRNEAKOPSUKCYGIPMLPTNTSTIKYLLVKLOASLK 60  
|||||  
85 TITLRRNNLVVMGVSDPFNGKCRHYHFNDDTSTERTDVENTLCSSSSRVAMSINYN 144  
|||||  
61 TITLRRNNLVVMGVSDPDY-NKCRHYHFNDDTSTERTDVENTLCPSSNPRVAKPDINYN 119  
|||||  
145 SLYPTMEKKAENSRNOVLGQIOLSSDICKISGVDSFFVKTAEFLLVAIQMVSAARE 204  
|||||  
120 GLYPTLEKKAGVTSRNEVLGQIOLSSDICKISGVDSFFVKTAEFLLVAIQMVSAARE 179  
|||||  
205 KYENOVKTNFRAPYDPKVINLEKWKIKSEAIHNKANGALPKPELVDAKGTWIVL 264  
|||||  
180 KYENOVKTNFRAPYDPKVINLEKWKIKSEAIHNKANGALPKPELVDAKGTWIVL 239  
|||||  
265 RVDEINROVALLKYNVGTCTT 286  
|||||  
240 RVDEINROVALLKYNVGTCTT 261

## RESULT 4

RIP2\_PHYAM STANDARD; PRT: 310 AA.  
ID RIP2\_PHYAM STANDARD; PRT: 310 AA.  
AC O40772;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Antiviral protein 2 precursor (PAP-II) (Ribosome-inactivating protein) (rRNA N-glycosidase) (EC 3.2.2.22).  
DE PAP2 OR PAP1.  
DE Phytolacca americana (Common pokeweed) (Virginian pokeweed).  
DE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
DE Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
DE Caryophyllidae; Caryophyllales; Phytolaccaceae; Phytolacca.  
DE NCBI\_TaxID=3527;  
DE [1]  
DE SEQUENCE FROM N.A.  
DE TISSUE=Leaf;  
DE MEDLINE=94307398; PubMed=6091760;  
DE Poyet J.-L., Radom J., Hoeveler A.;  
DE "Isolation and characterization of a cDNA clone encoding the pokeweed  
DE antiviral protein II from Phytolacca americana and its expression in  
DE E. coli";  
DE FEBS Lett. 347:268-272(1994).  
DE [2]  
DE SEQUENCE OF 26-55.  
DE TISSUE=Leaf;  
DE MEDLINE=85023392; PubMed=6091760;  
DE Bjorn M.J., Larrick J., Pistak M., Wilson K.J.;  
DE "Characterization of translational inhibitors from Phytolacca  
DE americana. Amino-terminal sequence determination and antibody-  
DE inhibitor conjugates";  
DE Biochim. Biophys. Acta 790:154-163(1984).  
DE !- FUNCTION: INHIBITS VIRAL INFECTION OF PLANTS. INHIBITS PROTEIN  
DE SYNTHESIS IN BOTH PROKARYOTES AND EUKARYOTES.  
DE !- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
DE specific adenosine on the 28S rRNA.  
DE !- TISSUE SPECIFICITY: EXPRESSED IN LATE SUMMER LEAVES.  
DE !- DEVELOPMENTAL STAGE: EXPRESSED PROGRESSIVELY WITH THE AGING OF THE  
DE PLANT.  
DE !- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
DE TYPE 1 RIP SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL Outstation  
at the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: X78628; CAA55342.1;  
DR HSP: Q03464; IAPA.  
DR InterPro: IPR001574; RIP.  
DR Pfam: PF00161; RIP: 1.  
DR PRINTS: PR00396; SHIGARICIN.  
DR PROSITE: PS00275; SHIGARICIN.  
KW Plant defense: Antiviral; Protein synthesis inhibitor; hydrolase;  
KW Toxin; Signal.  
FT SIGNAL. 1 25  
FT CHAIN 26 ?  
FT PROPEP 310  
FT ACT\_SITE 157 197  
FT DISULFID 57 284  
FT DISULFID 106 123  
SQ SEQUENCE 310 AA: 34694 MW: 40380010725909F CRC64;  
Query Match 31.6% Score 511; DB 1; Length 310;  
Best Local Similarity 39.9% Pred No. 1.7e-33;  
Matches 130; Conservative 53; Mismatches 111; Indels 32; Gaps 14;

QY 1 MKVMLVVVVTGLIAWLIAPTSTCAINTITFDAGNATINATYATMESLRNOAKDPKLCYVG 60  
DB 3 MKVLEVVGLAISIMLMITPP---ASSNIVDFDENATPETYSNELTSLREAVKDKKLTJNG 59  
QY 61 IPLPOTNSTPKYLLVKLGANLKTITLMLRRNNLYVMGVSDPFNGKCRHYHFNDDTST 120  
DB 60 MINATTLTEQPKYVILVDLFGS-GTFTLAIRGNLYIEGYSDIYNG-KCRYHFNKND--- 113  
QY 121 ERTDVENTLCSSSSR--VAMSYNYSLSYPTMEKKAENSRNOVLGQIOLSSDICKING 178  
DB 114 SESDAQETVCPGDKSKPGUTONNIPYEKSYKMGESKG--GARTKLGKGLTKLSRWCKING 171  
QY 179 VDSFP---VKTAEFLLVAIQMVSAAREFYENOVKTNFRA--FYDDPKVINLEKMW 232  
DB 172 KDATDOKQYKNEAEFLLVAIQMVSAAREFYENOVKTNFRA--FYDDPKVINLEKMW 231  
QY 233 GKISEAIHNKNG----ALPKLELVDAKGTWIVLRVDEINROV-ALUKYVNGTCTT 286  
DB 232 DSVSKVI--AKVGTSDUSTVITLPCDLKDNENKFWITATMNLKNOIMALLTHV--TCNV- 286  
QY 287 YONAMFSOVIIISTYNNYMSNLGLDFE 312  
DB 287 -XSSMPFETMSYYNTSISNLGE-FE 310

## RESULT 5

RIP6\_SAPOF STANDARD; PRT: 299 AA.  
ID RIP6\_SAPOF STANDARD; PRT: 299 AA.  
AC P20656; O41392;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1991 (Rel. 34, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Ribosome-inactivating protein saporin-6 precursor (SAP-6) (SU-6)  
DE (rRNA N-glycosidase) (EC 3.2.2.22).  
DE SAP6.  
DE Saponaria officinalis (Common soapwort).  
DE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
DE Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
DE Caryophyllidae; Caryophyllales; Caryophyllaceae; Saponaria.  
DE NCBI\_TaxID=3572;  
DE [1]  
DE SEQUENCE OF 1-283 FROM N.A., AND SEQUENCE OF 25-115; 206-213 AND  
DE 234-277.  
DE TISSUE=Leaf;  
DE MEDLINE=89338421; PubMed=2547612;  
DE Benatti L., Saccardo M.R., Dani M., Nitti G., Sassano M.,  
DE Lorenzetti R., Lippi D.A., Sorla M.;  
DE "Nucleotide sequence of cDNA coding for saporin-6, a type-1 ribosome-  
DE inactivating protein from Saponaria officinalis";  
DE Eur. J. Biochem. 183:463-470(1989).  
DE [2]  
DE SEQUENCE OF 25-277 FROM N.A.





[6] X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).  
MEDLINE=94016586; PubMed=8411176;  
Monzinger A.F., Collins E.J., Ernst S.R., Irvin J.D., Robertus J.D.:  
"The 2.5 A structure of pokeweed antiviral protein.";  
J. Mol. Biol. 233:705-715(1993).  
[7] X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 23-284.  
MEDLINE=99421320; PubMed=10493577;  
Kurinov I.V., Myers D.E., Irvin J.D., Uckun F.M.:  
"X-ray crystallographic analysis of the structural basis for the  
interactions of pokeweed antiviral protein with its active site  
inhibitor and ribosomal RNA substrate analogs.";  
Protein Sci. 8:1765-1772(1999).  
-!- FUNCTION: INHIBITS VIRAL INFECTION OF PLANTS, AND PROTEIN  
SYNTHESIS IN VITRO.  
-!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
specific adenosine on the 28S rRNA.  
-!- TISSUE SPECIFICITY: SPRING LEAVES AND ROOTS.  
-!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
TYPE 1 RIP SUBFAMILY.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation.  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

EMBL: X55383; CAA39054.1;  
PIR: S02792; S02792;  
PIR: S13469; S13469.  
PDB: 1PAF; 31-JAN-94.  
PDB: 1PAG; 31-JAN-94.  
PDB: 1QCI; 15-SEP-99.  
PDB: 1QCG; 15-SEP-99.  
PDB: 1QCI; 15-SEP-99.  
InterPro: IPR001574; RIP.  
Pfam: PF00161; RIP; 1.  
PRINTS: P000396; SHIGARICIN.  
PROSITE: PS00275; SHIGA\_RICIN; 1.  
KW Plant defense; Antiviral; Protein synthesis inhibitor; Hydrolase;  
Toxin; Signal; 3D-structure.  
SIGNAL 1 22  
CHAIN 23 285  
PROPEP 286 313  
ACT\_SITE 198 198  
DISULFID 56 281  
DISULFID 107 128  
STRAND 25 29  
HELIX 35 49  
STRAND 55 56  
TURN 57 58  
STRAND 59 61  
TURN 65 66  
STRAND 71 78  
HELIX 79 81  
STRAND 82 89  
TURN 90 93  
STRAND 94 102  
TURN 103 105  
STRAND 106 112  
TURN 113 114  
TURN 117 117  
TURN 118 127  
TURN 131 132  
STRAND 134 136  
STRAND 139 139  
HELIX 145 152  
TURN 153 153  
TURN 157 158  
STRAND 162 162

ANTIVIRAL PROTEIN I.

BY SIMILARITY.

RESULT 2

RIPA\_PHYAM

ID RIPA\_PHYAM STANDARD: PRT; 294 AA.  
AC Q03464;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-OCT-1993 (Rel. 27, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Antiviral protein alpha precursor (PAP-alpha) (Ribosome-inactivating  
protein) (rRNA N-glycosidase) (EC 3.2.2.22).  
OS Phytolacca americana (Common pokeweed) (Virginiaian pokeweed).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Caryophyllales; Caryophyllaceae; Phytolaccaceae; Phytolacca.  
OX NCBI\_TaxID=3527;  
RN (1)  
RP SEQUENCE FROM N.A.  
RC TISSUE=Seed, Leaf, and root;  
RX MEDLINE=93099240; PubMed=1281438;  
RA Kataoka J., Habuka N., Masuta C., Miyano M., Koizumi A.;  
RT "Isolation and analysis of a genomic clone encoding a pokeweed  
antiviral protein.";  
RL Plant Mol. Biol. 20:879-886(1992).  
RN (2)  
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).  
RX MEDLINE=95010127; PubMed=7925458;  
RA Ago H., Kataoka J., Tsuge H., Iwakura N., Inagaki E., Noma M.

FT HELIX 164 174  
FT TURN 175 176  
FT HELIX 182 195  
FT TURN 196 196  
FT HELIX 197 201  
FT HELIX 203 211  
FT TURN 212 214  
FT STRAND 217 217  
FT HELIX 221 239  
FT STRAND 241 241  
FT TURN 242 243  
FT STRAND 244 252  
FT TURN 254 255  
FT STRAND 258 263  
FT HELIX 264 270  
FT STRAND 274 274  
SQ SEQUENCE 313 AA; 35219 MW; 2C5782861EBA57F5 CRC64;

Query Match 76.4%; Score 1235.5; DB 1; Length 313;  
Best Local Similarity 75.9%; Pred. No. 4.9e-91;  
Matches 239; Conservative 33; Mismatches 40; Indels 3; Gaps 2;

QY 1 MKVMLVVVTLIAWLIAPTSTCAINTITFDAGNATINKYATEPESLNQAKDPKIKCYG 60  
DB 1 MKSMLVVTISI--WLIAPTSTMAVNTIYVGVSTTISKYATFLNDLRNEAKDPSIKCYG 58  
QY 61 IPMLPOTSTPKYLLVQLQGANLKITLMLRNLLYVMGYSDPFNGKCRVHIFNDIUST 120  
DB 59 IPMLPNTNPKYVLVLEQGSNKKITLMLRNLLYVMGYSDPFNPKCRVHIFNDISCT 118  
QY 121 ERTDVENTLCSSSSKVAHNSLYPTMKEKKAHVNSRNOVOLGIGIQICLSIDGKICWD 180  
DB 119 ERQDVETTLCPNANSRVSKNINFDSPRTLESKAGKVSQVQLGILDSNIGKISGM 178  
QY 181 SPVKTAEAFLLVAIQMVSEARFKYIENOVKTNFNAPYDPKVINLEEKWKISEAH 240  
DB 179 SFTEKTEAEFLVAIQMVSEARFKYIENOVKTNFNAPYDPKVINLEEKWKISEAH 238  
QY 241 NAKNGALPKLELVDAKGTWIVLRVDEINRDVALLKYVNGTCQTTY-QNAMFSQVILST 299  
DB 239 DAKNGVLPKPLELVDAKGTWIVLRVDEINRDVALLKYVNGTCQTTYQNAMEPOLINST 298  
QY 300 YYNYSNGLDLEEGF 314  
DB 299 YYNYSNGLDLEEGF 313

GenCore version 5.1.4\_p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

protein - protein search, using sw model

Run on: May 28, 2003, 09:46:19 : Search time 25 seconds  
(without alignments)  
520.943 Million cell updates/sec

Title: US-09-978-274A-2  
Perfect score: 1617  
Sequence: 1 MKVHLVVVTLIAIAAPT.....VLIITYYNSNLGDLFEGF 314

Scoring table:  
Gapop 10.0 , Gapext 0.5

searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Seq#	Score	Query Match	Length	ID	Description
1	1235.5	76.4	313	1 RIPL_PHYAM	P10297 phytolacca
2	1113.5	68.9	294	1 RIPA_PHYAM	Q03464 phytolacca
3	1096.5	67.8	261	1 RIPS_PHYAM	P23339 phytolacca
4	511	31.6	310	1 RIPL_PHYAM	Q40772 phytolacca
5	378.5	23.4	299	1 RIPA_PHYAM	P20656 saponaria o
6	372	23.0	292	1 RIPL_PHYAM	P27559 saponaria o
7	345.5	21.4	293	1 RIPO_PHYAM	P24476 dianthus ca
8	337	20.8	253	1 RIPS_PHYAM	Q41389 saponaria o
9	336	20.8	253	1 RIPL_PHYAM	Q41391 saponaria o
10	330	20.4	316	1 RIPA_PHYAM	P33186 gelonium mu
11	305.5	18.9	576	1 RIPA_PHYAM	P2879 ricinus com
12	300	18.6	294	1 RIPL_PHYAM	P56626 trichosanthe
13	285	17.6	236	1 RIPA_PHYAM	P27560 saponaria o
14	279.5	17.3	278	1 RIPA_PHYAM	P13326 marabalis j
15	271.5	16.8	564	1 RIPA_PHYAM	P68184 bryonia dio
16	263	16.3	282	1 RIPA_PHYAM	P8184 bryonia dio
17	253	15.6	286	1 RIPL_PHYAM	Q9184 cucumis fig
18	249.5	15.4	286	1 RIPL_PHYAM	P16094 momordica c
19	246	15.2	289	1 RIPL_PHYAM	P09989 trichosanthe
20	245.5	15.2	289	1 RIPA_PHYAM	P24478 trichosanthe
21	238.5	14.7	563	1 RIPA_PHYAM	P33183 sambucus ni
22	236.5	14.6	286	1 RIPL_PHYAM	P29339 momordica b
23	235	14.5	250	1 RIPL_PHYAM	P22851 luffa cylin
24	232.5	14.4	290	1 RIPL_PHYAM	P33185 bryonia dio
25	232.5	14.4	528	1 RIPA_PHYAM	P11140 abrus prec
26	231	14.3	527	1 RIPA_PHYAM	Q06077 abrus prec
27	227	14.0	562	1 RIPA_PHYAM	P28590 abrus prec
28	207	12.8	277	1 RIPL_PHYAM	Q00465 luffa cylin
29	202	12.5	157	1 RIPL_PHYAM	P27561 saponaria o
30	200	12.4	254	1 RIPL_PHYAM	P81446 viscum albu
31	131	8.1	32	1 RIPL_PHYAM	P34967 phytolacca
32	130.5	8.1	315	1 RIPL_PHYAM	P08026 bacterioph
33	127.5	7.9	315	1 RIPL_PHYAM	P10149 bacterioph

34	121	7.5	315	1 SLTA_00933	P09385 bacterioph
35	119	7.4	300	1 RIPL_PHYAM	P25891 ze mays (m
36	112.5	7.0	280	1 RIPL_PHYAM	P04399 hordeum vul
37	111.5	6.9	560	1 RIPL_PHYAM	Q00531 hordeum vul
38	111	6.9	3343	1 YOG7_00933	P34616 caenorhabdl
39	109.5	6.8	280	1 RIPL_PHYAM	P22244 hordeum vul
40	105.5	6.5	301	1 RIPL_PHYAM	P28522 ze mays (m
41	103	6.4	304	1 RIPL_PHYAM	P25892 ze mays (m
42	97.5	6.0	485	1 HXKA_YEAST	P04806 saccharomyc
43	94.5	5.8	385	1 GBAS_CABEL	Q20701 caenorhabdl
44	92.5	5.7	303	1 RIPL_PHYAM	P10593 ze mays (m
45	92.5	5.7	484	1 YA62_METJA	G58462 methanotoco

ALIGNMENTS

RESULT 1	RIPL_PHYAM	STANDARD:	PRT:	313 AA.
AC	P10297			
DT	01-MAR-1989 (Rel. 10, Created)			
DT	01-OCT-1993 (Rel. 27, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Antiviral protein I precursor (PAP-I) (PAP-C) (Ribosome-inactivating protein) (rRNA N-glycosidase) (EC 3.2.2.22).			
DE	PAP1			
GN	Phytolacca americana (Common pokeweed) (Virginian pokeweed).			
OS	Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:			
OC	Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots:			
OC	Caryophyllales: Caryophyllales: Phytolaccaceae: Phytolacca.			
OX	NCBI_TaxID=3527;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 23-62.			
RC	TISSUE=Leaf;			
RC	MEDLINE=92003676; PubMed=1912488;			
RA	Lin Q., Chen Z.C., Antoniw J.F., White R.F.;			
KT	"Isolation and characterization of a cDNA clone encoding the anti-viral protein from Phytolacca americana.";			
RT	Plant Mol. Biol. 17:609-614(1991).			
RL	[2]			
RN	SEQUENCE OF 23-65.			
RP	MEDLINE=89193489; PubMed=2930487;			
RA	Barbieri L., Bolognesi A., Ceini P., Falasca A.I., Minghetti A.,			
RA	Garofano L., Guicciardi A., Lepel D., Miller S.P.;			
RT	"Ribosome-inactivating proteins from plant cells in culture.";			
RL	Biochem. J. 257:801-807(1989).			
RC	[3]			
RN	SEQUENCE OF 23-54.			
RP	TISSUE=Leaf;			
RC	MEDLINE=83290867; PubMed=6885760;			
RA	Houston L.L., Ramakrishnan S., Hermodson M.A.;			
RT	"Seasonal variations in different forms of pokeweed antiviral protein, a potent inactivator of ribosomes.";			
RL	J. Biol. Chem. 258:9601-9604(1983).			
RN	[4]			
RP	SEQUENCE OF 23-54.			
RC	TISSUE=Leaf;			
RA	MEDLINE=85023392; PubMed=6091760;			
RA	Bjorn M.J., Larrick J., Platak M., Wilson K.J.;			
RT	"Characterization of translational inhibitors from Phytolacca americana: amino-terminal sequence determination and antibody-inhibitor conjugates.";			
RL	Biochim. Biophys. Acta 790:154-163(1984).			
RN	[5]			
RP	SEQUENCE OF 23-54.			
RC	TISSUE=Root;			
RA	MEDLINE=91064383; PubMed=2248976;			
RA	Bolognesi A., Barbieri L., Abbondanza A., Falasca A.I., Carnicelli D.,			
RA	Battelli M.G., Shirin F.;			
RT	"Purification and properties of new ribosome-inactivating proteins with RNA N-glycosidase activity.";			
RL	Biochim. Biophys. Acta 1087:273-302(1990).			

Query Match 2.5%; Score 8; DB 1; Length 328;  
Best Local Similarity 100.0%; Pred. No. 7.9;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

212 KTNFNRAF 219  
|||||||  
60 KTNFNRAF 67

## RESULT 14

D-alanine/glycine permease [imported] - Chlamydomophila pneumoniae (strain J138)  
C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 23-Mar-2001  
C:Accession: B86600  
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ise,  
Nucleic Acids Res. 28, 2311-2314, 2000  
A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.  
A:Reference number: A86491; MUID:20330349; PMID:10871362  
A:Accession: B86600  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-458 <STO>  
A:Cross-references: GB:BA000008; MID:98979250; PIDN:BAA99084.1; GSPDB:CN00142  
A:Experimental source: strain J138  
C:Genetics:  
A:Gene: daqA\_2  
C:Superfamily: sodium-dependent D-alanine/glycine transport protein

Query Match 2.5%; Score 8; DB 2; Length 458;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

270 NR0VALLK 277  
|||||||  
436 NR0VALLK 443

## RESULT 15

D-alanine/glycine permease - Chlamydomophila pneumoniae (strain CWL029)  
C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 05-May-2000  
C:Accession: B72024  
R:Kalman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;  
Nature Genet. 21, 385-389, 1999  
A:Title: Comparative genomes of chlamydia pneumoniae and C. trachomatis.  
A:Reference number: A72000; MUID:99206606; PMID:10192388  
A:Accession: B72024  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-458 <ARN>  
A:Cross-references: GB:AE001668; GB:AE001363; MID:94377192; PIDN:AAD19014.1; PID:9437719  
A:Experimental source: strain CWL029  
C:Genetics:  
A:Gene: daqA\_2  
C:Superfamily: sodium-dependent D-alanine/glycine transport protein

Query Match 2.5%; Score 8; DB 2; Length 458;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

270 NR0VALLK 277  
|||||||  
436 NR0VALLK 443

Search completed: May 28, 2003, 10:12:40  
Time: 47 secs

Printed 11/17/2004

Submitted to the Brookhaven Protein Data Bank, July 1994  
 Reference number: A67089; PDB:1MRH  
 Contents: annotation; X-ray crystallography, 2.0 angstroms, residues 24-77, 'R', 79-132.  
 Function:  
 Description: hydrolyzes the N-glycosidic bond of a specific adenosine in 28S rRNA that  
 Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
 Keywords: glycoprotein; glycosidase; hydrolase; lectin; seed; toxin  
 F:1-18/Domain: signal sequence #status predicted <SIG>  
 F:19-23/Domain: amino-terminal propeptide #status predicted <PRO>  
 F:24-269/Product: rRNA N-glycosidase alpha-momorcharin #status experimental <MAT>  
 F:27-266/Domain: rRNA N-glycosidase homology <RNG>  
 F:270-286/Domain: carboxyl-terminal propeptide #status predicted <CTP>  
 F:293,183,186/Active site: Tyr, Glu, Arg #status predicted  
 F:250/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 2.9% Score 9; DB 2; Length 286;  
 Best Local Similarity 100.0%; Pred. No. 0.7;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 200 EAARFKYIE 208  
 |||||  
 DB 183 EAARFKYIE 191

RESULT 10  
 S25560  
 rRNA N-glycosidase (EC 3.2.2.22) momordin II - balsam apple  
 C:Species: Momordica balsamina (balsam apple)  
 C:Date: 25-Feb-1994 #sequence\_revision 03-Aug-1995 #text\_change 20-Aug-1999  
 C:Accession: S25560  
 R:Ortigao, M.; Better, M.  
 Nucleic Acids Res. 20, 4662, 1992

Title: Momordin II, a ribosome inactivating protein from Momordica balsamina, is homologous to the N-glycosidase from Momordica charantia  
 Reference number: S25560; MUID: 93027170; PMID:1408771  
 Accession: S25560  
 Status: preliminary  
 Molecule type: mRNA  
 Residues: 1-286 <ORT>

C:Cross-references: EMBL:212175; NID:919525; PDB:CAA78166.1; PTD:919526  
 C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
 C:Keywords: glycosidase; hydrolase  
 F:27-264/Domain: rRNA N-glycosidase homology <RNG>

Query Match 2.9% Score 9; DB 2; Length 286;  
 Best Local Similarity 100.0%; Pred. No. 0.7;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 200 EAARFKYIE 208  
 |||||  
 DB 181 EAARFKYIE 189

RESULT 11  
 T32320  
 hypothetical protein C41H7.6 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 21-Jul-2000  
 C:Accession: T32320  
 R:Geisel, C.; Wamsley, P.; Elliott, G.; Smith, A.  
 Submitted to the EMBL Data Library, September 1997  
 Description: The sequence of C. elegans cosmid C41H7.

Reference number: 221150  
 Accession: T32320  
 Status: preliminary; translated from GB/EMBL/DDBJ  
 Molecule type: DNA  
 Residues: 1-321 <GEI>  
 C:Cross-references: EMBL:AF025450; PDB:AA870937.1; GSPDB:GN00020; CESP:C41H7.6  
 C:Experimental source: strain Bristol N2; clone C41H7  
 C:Genetics:  
 A:Gene: CESP:C41H7.6  
 A:Map position: 2  
 A:Introns: 93/1; 183/2  
 C:Superfamily: Caenorhabditis elegans hypothetical protein C08A9.6

Query Match 2.5% Score 8; DB 2; Length 321;  
 Best Local Similarity 100.0%; Pred. No. 7.8;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 305 SNLGDLEF 312  
 |||||  
 DB 294 SNLGDLEF 301

RESULT 12  
 T32321  
 hypothetical protein C41H7.5 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 21-Jul-2000  
 C:Accession: T32321  
 R:Geisel, C.; Wamsley, P.; Elliott, G.; Smith, A.  
 Submitted to the EMBL Data Library, September 1997  
 Description: The sequence of C. elegans cosmid C41H7.  
 Reference number: 221150  
 Accession: T32321  
 Status: preliminary; translated from GB/EMBL/DDBJ  
 Molecule type: DNA  
 Residues: 1-327 <GEI>  
 C:Cross-references: EMBL:AF025450; PDB:AA870936.1; GSPDB:GN00020; CESP:C41H7.5  
 C:Experimental source: strain Bristol N2; clone C41H7  
 C:Genetics:  
 A:Gene: CESP:C41H7.5  
 A:Map position: 2  
 A:Introns: 99/1; 189/2  
 C:Superfamily: Caenorhabditis elegans hypothetical protein C08A9.6

Query Match 2.5% Score 8; DB 2; Length 327;  
 Best Local Similarity 100.0%; Pred. No. 7.9;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 305 SNLGDLEF 312  
 |||||  
 DB 300 SNLGDLEF 307

RESULT 13  
 C70039  
 two-component sensor histidine kinase homolog yvfr - Bacillus subtilis  
 C:Species: Bacillus subtilis  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 21-Jul-2000  
 C:Accession: C70039  
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azavedo, V.; Ber  
 C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capriano, V.; Carter, N.M.;  
 A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,  
 Nature 390, 249-256, 1997  
 Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gal  
 iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M  
 Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino  
 A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau  
 Y. M.; Ogawa, K.; Ogawara, A.; Oudea, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Porcete  
 Rieger, M.; Rivolta, C.; Kocha, E.; Roche, B.; Rose, M.; Sadaka, Y.; Sato, T.; Scani  
 A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se  
 akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpstra, P.; Toyonari, A.; Tsubota, V.; Uchiya  
 T.; Winters, P.; Wipal, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida  
 A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Zanchin, A.  
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis  
 A:Reference number: A69580; MUID:98044033; PMID:9384377  
 Accession: C70039  
 Status: preliminary; nucleic acid sequence not shown; translation not shown  
 Molecule type: DNA  
 Residues: 1-328 <KUN>  
 C:Cross-references: GB:299121; GB:AL009126; NID:92635827; PDB:CA815412.1; PTD:926359  
 C:Experimental source: strain 168  
 C:Genetics:  
 A:Gene: yvfr  
 C:Superfamily: probable Bacillus subtilis two-component sensor histidine kinase yocF



Query Match 9.6%; Score 30; DB 2; Length 313;  
Best Local Similarity 100.0%; Pred. No. 5.9e-22;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 190 FLVAIQMVSEARFKYIENQVKTNFNR 219  
Db 188 FLVAIQMVSEARFKYIENQVKTNFNR 217  
A: Accession: JE0401  
A: Molecule type: protein  
A: Residues: 1-261 <NON>  
A: Experimental source: seed  
C: Comment: This protein prevents the replication of a number of plant viruses, and inhibits the replication of the tobacco etch virus.  
C: Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
C: Keywords: disulfide bond; glycoprotein  
F: 6-254/Domain: rRNA N-glycosidase homology <RNG>  
F: 10,44,255/Binding site: Carboxylate (Asn) (covalent) #status experimental  
F: 34-258,84-105/Disulfide bonds: #status experimental

Query Match 8.9%; Score 28; DB 2; Length 261;  
Best Local Similarity 100.0%; Pred. No. 5.2e-20;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 190 FLVAIQMVSEARFKYIENQVKTNFNR 217  
Db 165 FLVAIQMVSEARFKYIENQVKTNFNR 192  
A: Accession: JE0401  
A: Molecule type: protein  
A: Residues: 1-261 <NON>  
A: Experimental source: seed  
C: Comment: This protein prevents the replication of a number of plant viruses, and inhibits the replication of the tobacco etch virus.  
C: Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
C: Keywords: disulfide bond; glycoprotein  
F: 6-254/Domain: rRNA N-glycosidase homology <RNG>  
F: 10,44,255/Binding site: Carboxylate (Asn) (covalent) #status experimental  
F: 34-258,84-105/Disulfide bonds: #status experimental

Query Match 8.9%; Score 28; DB 2; Length 261;  
Best Local Similarity 100.0%; Pred. No. 5.2e-20;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 190 FLVAIQMVSEARFKYIENQVKTNFNR 217  
Db 165 FLVAIQMVSEARFKYIENQVKTNFNR 192  
A: Accession: JE0401  
A: Molecule type: protein  
A: Residues: 1-261 <NON>  
A: Experimental source: seed  
C: Comment: This protein prevents the replication of a number of plant viruses, and inhibits the replication of the tobacco etch virus.  
C: Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
C: Keywords: disulfide bond; glycoprotein  
F: 6-254/Domain: rRNA N-glycosidase homology <RNG>  
F: 10,44,255/Binding site: Carboxylate (Asn) (covalent) #status experimental  
F: 34-258,84-105/Disulfide bonds: #status experimental

Query Match 8.9%; Score 28; DB 2; Length 261;  
Best Local Similarity 100.0%; Pred. No. 5.2e-20;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 190 FLVAIQMVSEARFKYIENQVKTNFNR 217  
Db 165 FLVAIQMVSEARFKYIENQVKTNFNR 192  
A: Accession: JE0401  
A: Molecule type: protein  
A: Residues: 1-261 <NON>  
A: Experimental source: seed  
C: Comment: This protein prevents the replication of a number of plant viruses, and inhibits the replication of the tobacco etch virus.  
C: Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
C: Keywords: disulfide bond; glycoprotein  
F: 6-254/Domain: rRNA N-glycosidase homology <RNG>  
F: 10,44,255/Binding site: Carboxylate (Asn) (covalent) #status experimental  
F: 34-258,84-105/Disulfide bonds: #status experimental

Query Match 9.6%; Score 30; DB 2; Length 313;  
Best Local Similarity 100.0%; Pred. No. 5.9e-22;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 190 FLVAIQMVSEARFKYIENQVKTNFNR 219  
Db 188 FLVAIQMVSEARFKYIENQVKTNFNR 217  
A: Accession: JE0401  
A: Molecule type: protein  
A: Residues: 1-261 <NON>  
A: Experimental source: seed  
C: Comment: This protein prevents the replication of a number of plant viruses, and inhibits the replication of the tobacco etch virus.  
C: Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
C: Keywords: disulfide bond; glycoprotein  
F: 6-254/Domain: rRNA N-glycosidase homology <RNG>  
F: 10,44,255/Binding site: Carboxylate (Asn) (covalent) #status experimental  
F: 34-258,84-105/Disulfide bonds: #status experimental

Query Match 8.9%; Score 28; DB 2; Length 261;  
Best Local Similarity 100.0%; Pred. No. 5.2e-20;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 190 FLVAIQMVSEARFKYIENQVKTNFNR 219  
Db 188 FLVAIQMVSEARFKYIENQVKTNFNR 217  
A: Accession: JE0401  
A: Molecule type: protein  
A: Residues: 1-261 <NON>  
A: Experimental source: seed  
C: Comment: This protein prevents the replication of a number of plant viruses, and inhibits the replication of the tobacco etch virus.  
C: Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
C: Keywords: disulfide bond; glycoprotein  
F: 6-254/Domain: rRNA N-glycosidase homology <RNG>  
F: 10,44,255/Binding site: Carboxylate (Asn) (covalent) #status experimental  
F: 34-258,84-105/Disulfide bonds: #status experimental

Query Match 8.9%; Score 28; DB 2; Length 261;  
Best Local Similarity 100.0%; Pred. No. 5.2e-20;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 190 FLVAIQMVSEARFKYIENQVKTNFNR 217  
Db 165 FLVAIQMVSEARFKYIENQVKTNFNR 192  
A: Accession: JE0401  
A: Molecule type: protein  
A: Residues: 1-261 <NON>  
A: Experimental source: seed  
C: Comment: This protein prevents the replication of a number of plant viruses, and inhibits the replication of the tobacco etch virus.  
C: Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
C: Keywords: disulfide bond; glycoprotein  
F: 6-254/Domain: rRNA N-glycosidase homology <RNG>  
F: 10,44,255/Binding site: Carboxylate (Asn) (covalent) #status experimental  
F: 34-258,84-105/Disulfide bonds: #status experimental

Query Match 8.9%; Score 28; DB 2; Length 261;  
Best Local Similarity 100.0%; Pred. No. 5.2e-20;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 190 FLVAIQMVSEARFKYIENQVKTNFNR 217  
Db 165 FLVAIQMVSEARFKYIENQVKTNFNR 192  
A: Accession: JE0401  
A: Molecule type: protein  
A: Residues: 1-261 <NON>  
A: Experimental source: seed  
C: Comment: This protein prevents the replication of a number of plant viruses, and inhibits the replication of the tobacco etch virus.  
C: Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
C: Keywords: disulfide bond; glycoprotein  
F: 6-254/Domain: rRNA N-glycosidase homology <RNG>  
F: 10,44,255/Binding site: Carboxylate (Asn) (covalent) #status experimental  
F: 34-258,84-105/Disulfide bonds: #status experimental

R: Kataoka, J.; Miyano, M.; Habuka, N.; Masura, C.; Kolwal, A.  
Submitted to the EMBL Data Library, February 1992  
A: Reference number: S36928  
A: Accession: S36928  
A: Molecule type: DNA  
A: Residues: 1-278 <KAT>  
A: Cross-references: EMBL: D10569; NID: g217953; PIDN: BAA01425.1; PIR: Q217454  
R: Kataoka, J.; Habuka, N.; Furuno, M.; Miyano, M.; Takanami, Y.; Kolwal, A.  
J. Biol. Chem. 266, 8426-8430, 1991  
A: Title: DNA sequence of Mirabilis antiviral protein (MAP), a ribosome-inactivating protein  
A: Reference number: A39817; MUID: 91217082; PMID: 2022657  
A: Accession: A39817  
A: Molecule type: mRNA  
A: Residues: 1-278 <KAT>  
A: Cross-references: GB: D90347  
R: Habuka, N.; Murakami, Y.; Noma, M.; Kudo, T.; Horikoshi, K.  
J. Biol. Chem. 264, 6629-6637, 1989  
A: Title: Amino acid sequence of Mirabilis antiviral protein, total synthesis of its gene  
A: Reference number: A33871; MUID: 89214063; PMID: 2708328  
A: Accession: A33871  
A: Molecule type: protein  
A: Residues: 29-278 <HAB>  
R: Ago, H.; Kataoka, J.; Tsuge, K.; Habuka, N.; Inagaki, E.; Noma, M.; Miyano, M.  
Eur. J. Biochem. 225, 369-374, 1994  
A: Title: X-ray structure of a pokeweed antiviral protein, coded by a new genomic clone  
A: Reference number: S48718; MUID: 95010127; PMID: 7925458  
A: Accession: S48718  
A: Molecule type: protein  
A: Residues: 29-91, 'K', '93-99, 'K', '100-154, 'K', '156-160, 'K', '162-217, 'K', '218-249, 'L', '251-255, 'L'  
A: Genetics: 206/3  
C: Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
C: Keywords: glycosidase; hydrolase  
F: 29-278/Product: antiviral protein MAP #status experimental <MAT>  
F: 38-278/Domain: rRNA N-glycosidase homology <RNG>

Query Match 4.5%; Score 14; DB 2; Length 278;  
Best Local Similarity 100.0%; Pred. No. 6.5e-06;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 194 AIQMVSEARFKYI 207  
Db 190 AIQMVSEARFKYI 203  
A: Accession: S38528  
A: Molecule type: protein  
A: Residues: 1-278 <KAT>  
A: Cross-references: GB: D90347  
R: Habuka, N.; Murakami, Y.; Noma, M.; Kudo, T.; Horikoshi, K.  
J. Biol. Chem. 264, 6629-6637, 1989  
A: Title: Amino acid sequence of Mirabilis antiviral protein, total synthesis of its gene  
A: Reference number: A33871; MUID: 89214063; PMID: 2708328  
A: Accession: A33871  
A: Molecule type: protein  
A: Residues: 29-278 <HAB>  
R: Ago, H.; Kataoka, J.; Tsuge, K.; Habuka, N.; Inagaki, E.; Noma, M.; Miyano, M.  
Eur. J. Biochem. 225, 369-374, 1994  
A: Title: X-ray structure of a pokeweed antiviral protein, coded by a new genomic clone  
A: Reference number: S48718; MUID: 95010127; PMID: 7925458  
A: Accession: S48718  
A: Molecule type: protein  
A: Residues: 29-91, 'K', '93-99, 'K', '100-154, 'K', '156-160, 'K', '162-217, 'K', '218-249, 'L', '251-255, 'L'  
A: Genetics: 206/3  
C: Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
C: Keywords: glycosidase; hydrolase  
F: 29-278/Product: antiviral protein MAP #status experimental <MAT>  
F: 38-278/Domain: rRNA N-glycosidase homology <RNG>

Query Match 4.5%; Score 14; DB 2; Length 278;  
Best Local Similarity 100.0%; Pred. No. 6.5e-06;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 194 AIQMVSEARFKYI 207  
Db 190 AIQMVSEARFKYI 203  
A: Accession: S38528  
A: Molecule type: protein  
A: Residues: 1-278 <KAT>  
A: Cross-references: GB: D90347  
R: Habuka, N.; Murakami, Y.; Noma, M.; Kudo, T.; Horikoshi, K.  
J. Biol. Chem. 264, 6629-6637, 1989  
A: Title: Amino acid sequence of Mirabilis antiviral protein, total synthesis of its gene  
A: Reference number: A33871; MUID: 89214063; PMID: 2708328  
A: Accession: A33871  
A: Molecule type: protein  
A: Residues: 29-278 <HAB>  
R: Ago, H.; Kataoka, J.; Tsuge, K.; Habuka, N.; Inagaki, E.; Noma, M.; Miyano, M.  
Eur. J. Biochem. 225, 369-374, 1994  
A: Title: X-ray structure of a pokeweed antiviral protein, coded by a new genomic clone  
A: Reference number: S48718; MUID: 95010127; PMID: 7925458  
A: Accession: S48718  
A: Molecule type: protein  
A: Residues: 29-91, 'K', '93-99, 'K', '100-154, 'K', '156-160, 'K', '162-217, 'K', '218-249, 'L', '251-255, 'L'  
A: Genetics: 206/3  
C: Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
C: Keywords: glycosidase; hydrolase  
F: 29-278/Product: antiviral protein MAP #status experimental <MAT>  
F: 38-278/Domain: rRNA N-glycosidase homology <RNG>

Query Match 3.2%; Score 10; DB 2; Length 32;  
Best Local Similarity 100.0%; Pred. No. 0.011;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 ESLRNOAKDP 54  
Db 21 ESLRNOAKDP 30  
A: Accession: S38528  
A: Molecule type: protein  
A: Residues: 1-32 <PAR>  
A: Cross-references: GB: D90347  
R: Habuka, N.; Murakami, Y.; Noma, M.; Kudo, T.; Horikoshi, K.  
J. Biol. Chem. 264, 6629-6637, 1989  
A: Title: Amino acid sequence of Mirabilis antiviral protein, total synthesis of its gene  
A: Reference number: A33871; MUID: 89214063; PMID: 2708328  
A: Accession: A33871  
A: Molecule type: protein  
A: Residues: 29-278 <HAB>  
R: Ago, H.; Kataoka, J.; Tsuge, K.; Habuka, N.; Inagaki, E.; Noma, M.; Miyano, M.  
Eur. J. Biochem. 225, 369-374, 1994  
A: Title: X-ray structure of a pokeweed antiviral protein, coded by a new genomic clone  
A: Reference number: S48718; MUID: 95010127; PMID: 7925458  
A: Accession: S48718  
A: Molecule type: protein  
A: Residues: 29-91, 'K', '93-99, 'K', '100-154, 'K', '156-160, 'K', '162-217, 'K', '218-249, 'L', '251-255, 'L'  
A: Genetics: 206/3  
C: Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
C: Keywords: glycosidase; hydrolase  
F: 29-278/Product: antiviral protein MAP #status experimental <MAT>  
F: 38-278/Domain: rRNA N-glycosidase homology <RNG>





processing, the protein contains a 22 amino acid N-terminal signal sequence and a 29 amino acid C-terminal extension. PAP is a Type I ribosomal inhibitor protein (RIP) that catalyses the removal of a specific adenine residue from a highly conserved stem-loop structure in the 28S rRNA of eukaryotic ribosomes and also interferes with elongation factor-2 binding thus blocking protein translation. The PAP mutants can confer broad spectrum virus and fungus resistance to plants without causing plant cell death or sterility. They can also confer increased resistance to insects, bacteria and nematodes in plants. N.B. This sequence is not given in the specification but is generated from the wild type sequence reproduced in the specification.

Query Match 9.6%; Score 30; DB 18; Length 207;  
Best Local Similarity 100.0%; Pred. No. 2e-20;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

190 FLVATQMVSEARFKYIENQVKTNFRAF 219  
166 FLVATQMVSEARFKYIENQVKTNFRAF 195

RESULT 15

AAW14193  
O AAW14193 standard: protein: 208 AA.

AAW14193;

25-JUN-1997 (first entry)

Mature pokeweed antiviral protein deletion variant (1-208).

Pokeweed antiviral protein: mutant: phytotoxicity; anti-viral; virus: anti-fungal; fungus; plant; ribosomal inhibitor protein; adenine; insect; stem-loop; 28S rRNA; eukaryotic; ribosome; elongation factor-2 binding; translation; broad spectrum; resistance; cell death; sterility; nematode; bacterium.

Phytolacca americana.

W09703183-A1.

30-JAN-1997.

11-JUL-1996; 96WO-US11546.

11-JUL-1995; 95US-0500694.

11-JUL-1995; 95US-0500611.

(RUTF ) UNIV RUTGERS STATE NEW JERSEY.

Tumor NE;

MPI: 1997-119040/11.

New pokeweed antiviral protein mutants - which exhibit anti-viral and anti-fungal activity in plants and have reduced phytotoxicity

Claim 8: -; 64pp; English.

Proteins AAW14163-W14256 represent new pokeweed antiviral protein (PAP) mutants having reduced phytotoxicity but retaining anti-viral or anti-fungal activity in plants. The sequences' numbering corresponds to the 262 amino acid mature protein. The proteins AAW14165-W14217 represent serial deletions from the C-terminus whereas sequences AAW14218-55 are serial deletions from the N-terminus. Prior to processing, the protein contains a 22 amino acid N-terminal signal sequence and a 29 amino acid C-terminal extension. PAP is a Type I ribosomal inhibitor protein (RIP) that catalyses the removal of a specific adenine residue from a highly conserved stem-loop structure in the 28S rRNA of eukaryotic ribosomes and also interferes with elongation factor-2 binding thus blocking protein translation. The PAP mutants can

CC confer broad spectrum virus and fungus resistance to plants without causing plant cell death or sterility. They can also confer increased resistance to insects, bacteria and nematodes in plants. CC N.B. This sequence is not given in the specification but is generated from the wild type sequence reproduced in the specification.

XX Sequence 208 AA;

Query Match 9.6%; Score 30; DB 18; Length 208;  
Best Local Similarity 100.0%; Pred. No. 2e-20;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 190 FLVATQMVSEARFKYIENQVKTNFRAF 219  
DB 166 FLVATQMVSEARFKYIENQVKTNFRAF 195

Search completed: May 28, 2003, 10:09:42  
Job time : 78 secs

New pokeweed antiviral protein mutants - which exhibit anti-viral and anti-fungal activity in plants and have reduced phytotoxicity

Claim 8: - 64pp; English.

Proteins AAW14163-W14256 represent new pokeweed antiviral protein (PAP) mutants having reduced phytotoxicity but retaining anti-viral or anti-fungal activity in plants. The sequences' numbering corresponds to the 262 amino acid mature protein. The proteins AAW14165-W14217 represent serial deletions from the C-terminus whereas sequences AAW14218-55 are serial deletions from the N-terminus. Prior to processing, the protein contains a 22 amino acid N-terminal signal sequence and a 29 amino acid C-terminal extension. PAP is a type I ribosomal inhibitor protein (RIP) that catalyses the removal of a specific adenine residue from a highly conserved stem-loop structure in the 28S rRNA of eukaryotic ribosomes and also interferes with elongation factor-2 binding thus blocking protein translation. The PAP mutants can confer broad spectrum virus and fungus resistance to plants without causing plant cell death or sterility. They can also confer increased resistance to insects, bacteria and nematodes in plants. N.B. This sequence is not given in the specification but is generated from the wild type sequence reproduced in the specification.

Sequence 205 AA:  
Query Match 9.6%; Score 30; DB 18; Length 205;  
Best Local Similarity 100.0%; Pred. No. 2e-20;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

190 FLVATQMVSEAAARFYIENQVKTNFRAF 219  
166 FLVATQMVSEAAARFYIENQVKTNFRAF 195

AAW14195;  
25-JUN-1997 (first entry)  
Mature pokeweed antiviral protein deletion variant (1-206).

Pokeweed antiviral protein; mutant: phytotoxicity; anti-viral; virus; anti-fungal; fungus; plant; ribosomal inhibitor protein; adenine; insect; stem-loop; 28S rRNA; eukaryotic; ribosome; elongation factor-2 binding; translation; broad spectrum; resistance; cell death; sterility; nematode; bacterium.

Phytolacca americana.

WO9703183-A1.

30-JAN-1997.

11-JUL-1996; 96WO-US11546.

11-JUL-1995; 95US-0500694.

11-JUL-1995; 95US-0500611.

(RUTE ) UNIV RUTGERS STATE NEW JERSEY.

Turner NE;

WPI: 1997-119040/11.

New pokeweed antiviral protein mutants - which exhibit anti-viral and anti-fungal activity in plants and have reduced phytotoxicity

Claim 8: - 64pp; English.

Proteins AAW14163-W14256 represent new pokeweed antiviral protein (PAP) mutants having reduced phytotoxicity but retaining anti-viral or anti-fungal activity in plants. The sequences' numbering corresponds to the 262 amino acid mature protein. The proteins AAW14165-W14217 represent serial deletions from the C-terminus whereas sequences AAW14218-55 are serial deletions from the N-terminus. Prior to processing, the protein contains a 22 amino acid N-terminal signal sequence and a 29 amino acid C-terminal extension. PAP is a type I ribosomal inhibitor protein (RIP) that catalyses the removal of a specific adenine residue from a highly conserved stem-loop structure in the 28S rRNA of eukaryotic ribosomes and also interferes with elongation factor-2 binding thus blocking protein translation. The PAP mutants can confer broad spectrum virus and fungus resistance to plants without causing plant cell death or sterility. They can also confer increased resistance to insects, bacteria and nematodes in plants. N.B. This sequence is not given in the specification but is generated from the wild type sequence reproduced in the specification.

Sequence 206 AA:

Query Match 9.6%; Score 30; DB 18; Length 206;  
Best Local Similarity 100.0%; Pred. No. 2e-20;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

190 FLVATQMVSEAAARFYIENQVKTNFRAF 219  
166 FLVATQMVSEAAARFYIENQVKTNFRAF 195

AAW14194;  
25-JUN-1997 (first entry)  
Mature pokeweed antiviral protein deletion variant (1-207).

Pokeweed antiviral protein; mutant: phytotoxicity; anti-viral; virus; anti-fungal; fungus; plant; ribosomal inhibitor protein; adenine; insect; stem-loop; 28S rRNA; eukaryotic; ribosome; elongation factor-2 binding; translation; broad spectrum; resistance; cell death; sterility; nematode; bacterium.

Phytolacca americana.

WO9703183-A1.

30-JAN-1997.

11-JUL-1996; 96WO-US11546.

11-JUL-1995; 95US-0500694.

11-JUL-1995; 95US-0500611.

(RUTE ) UNIV RUTGERS STATE NEW JERSEY.

Turner NE;

WPI: 1997-119040/11.

New pokeweed antiviral protein mutants - which exhibit anti-viral and anti-fungal activity in plants and have reduced phytotoxicity

Claim 8: - 64pp; English.

Proteins AAW14163-W14256 represent new pokeweed antiviral protein (PAP) mutants having reduced phytotoxicity but retaining anti-viral or anti-fungal activity in plants. The sequences' numbering corresponds to the 262 amino acid mature protein. The proteins AAW14165-W14217 represent serial deletions from the C-terminus whereas sequences AAW14218-55 are serial deletions from the N-terminus. Prior to

```

30-JAN-1997.
11-JUL-1996; 96WO-US11546.
11-JUL-1995; 95US-0500694.
11-JUL-1995; 95US-0500611.
(RUTF ) UNIV RUTGERS STATE NEW JERSEY.
Tuner NE:
WPI: 1997-119040/11.
New pokeweed antiviral protein mutants - which exhibit anti-viral
and anti-fungal activity in plants and have reduced phytotoxicity
Claim 8: -: 64pp: English.
Proteins AAW14163-W14256 represent new pokeweed antiviral protein (PAP)
mutants having reduced phytotoxicity but retaining anti-viral or
anti-fungal activity in plants. The sequences, numbering corresponds to
the 262 amino acid mature protein. The proteins AAW14165-W14217
represent serial deletions from the C-terminus whereas sequences
AAW14218-55 are serial deletions from the N-terminus. Prior to
processing, the protein contains a 22 amino acid N-terminal signal
sequence and a 29 amino acid C-terminal extension. PAP is a type I
ribosomal inhibitor protein (RIP) that catalyses the removal of a
specific adenine residue from a highly conserved stem-loop structure in
the 28S rRNA of eukaryotic ribosomes and also interferes with elongation
factor-2 binding thus blocking protein translation. The PAP mutants can
confer broad spectrum virus and fungus resistance to plants without
causing plant cell death or sterility. They can also confer increased
resistance to insects, bacteria and nematodes in plants.
N.B. This sequence is not given in the specification but is generated
from the wild type sequence reproduced in the specification.
Sequence 203 AA:
Query Match 9.6%; Score 30; DB 18; Length 203;
Best Local Similarity 100.0%; Pred. No. 2e-20;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 190 FLVAIQMVSEAAAREKYIENQVKTNFRAF 219
DDB 166 FLVAIQMVSEAAAREKYIENQVKTNFRAF 195
RESULT 11
AAW14197
AAW14197 standard; protein; 204 AA.
AAW14197;
25-JUN-1997 (first entry)
Mature pokeweed antiviral protein deletion variant (1-204).
Pokeweed antiviral protein; mutant; phytotoxicity; anti-viral; virus;
anti-fungal; fungus; plant; ribosomal inhibitor protein; adenine; insect;
stem-loop; 28S rRNA; eukaryotic; ribosome; elongation factor-2 binding;
translation; broad spectrum; resistance; cell death; sterility; nematode;
bacterium.
Phytolacca americana.
WO9703183-A1.
30-JAN-1997.
11-JUL-1996; 96WO-US11546.
11-JUL-1995; 95US-0500694.
11-JUL-1995; 95US-0500611.
(RUTF ) UNIV RUTGERS STATE NEW JERSEY.
Tuner NE:
WPI: 1997-119040/11.
New pokeweed antiviral protein mutants - which exhibit anti-viral
and anti-fungal activity in plants and have reduced phytotoxicity
Claim 8: -: 64pp: English.
Proteins AAW14163-W14256 represent new pokeweed antiviral protein (PAP)
mutants having reduced phytotoxicity but retaining anti-viral or
anti-fungal activity in plants. The sequences, numbering corresponds to
the 262 amino acid mature protein. The proteins AAW14165-W14217
represent serial deletions from the C-terminus whereas sequences
AAW14218-55 are serial deletions from the N-terminus. Prior to
processing, the protein contains a 22 amino acid N-terminal signal
sequence and a 29 amino acid C-terminal extension. PAP is a type I
ribosomal inhibitor protein (RIP) that catalyses the removal of a
specific adenine residue from a highly conserved stem-loop structure in
the 28S rRNA of eukaryotic ribosomes and also interferes with elongation
factor-2 binding thus blocking protein translation. The PAP mutants can
confer broad spectrum virus and fungus resistance to plants without
causing plant cell death or sterility. They can also confer increased
resistance to insects, bacteria and nematodes in plants.
N.B. This sequence is not given in the specification but is generated
from the wild type sequence reproduced in the specification.
Sequence 204 AA:
Query Match 9.6%; Score 30; DB 18; Length 204;
Best Local Similarity 100.0%; Pred. No. 2e-20;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 190 FLVAIQMVSEAAAREKYIENQVKTNFRAF 219
DDB 166 FLVAIQMVSEAAAREKYIENQVKTNFRAF 195
RESULT 12
AAW14196
AAW14196 standard; protein; 205 AA.
AAW14196;
25-JUN-1997 (first entry)
Mature pokeweed antiviral protein deletion variant (1-205).
Pokeweed antiviral protein; mutant; phytotoxicity; anti-viral; virus;
anti-fungal; fungus; plant; ribosomal inhibitor protein; adenine; insect;
stem-loop; 28S rRNA; eukaryotic; ribosome; elongation factor-2 binding;
translation; broad spectrum; resistance; cell death; sterility; nematode;
bacterium.
Phytolacca americana.
WO9703183-A1.
30-JAN-1997.
11-JUL-1996; 96WO-US11546.
11-JUL-1995; 95US-0500694.
11-JUL-1995; 95US-0500611.
(RUTF ) UNIV RUTGERS STATE NEW JERSEY.
Tuner NE:
WPI: 1997-119040/11.

```



166 FLVIAQMVSAAAEKFKYIENQVKTNFNRAF 195

RESULT 6  
AAW14202

ID AAW14202 standard; protein: 199 AA.

XX

XX

AC

AC

25-JUN-1997 (first entry)

XX

XX

DE

Mature pokeweed antiviral protein deletion variant (1-199).

XX

XX

OS

Phytolacca americana.

XX

XX

PN

W09703183-Al.

XX

XX

PD

30-JAN-1997.

XX

XX

PF

11-JUL-1996; 96WO-US11546.

XX

XX

PP

11-JUL-1995; 95US-0500694.

XX

XX

PR

11-JUL-1995; 95US-0500611.

XX

XX

PA

(RUTF ) UNIV RUTGERS STATE NEW JERSEY.

XX

XX

PT

New pokeweed antiviral protein mutants - which exhibit anti-viral

and anti-fungal activity in plants and have reduced phytotoxicity

XX

XX

PS

Tumor NE;

XX

XX

WPI: 1997-119040/11.

XX

XX

CC

Proteins AAW14163-W14256 represent new pokeweed antiviral protein (PAP)

mutants having reduced phytotoxicity but retaining anti-viral or

anti-fungal activity in plants. The sequences' numbering corresponds to

the 262 amino acid mature protein. The proteins AAW14163-W14217

166 FLVIAQMVSAAAEKFKYIENQVKTNFNRAF 195

RESULT 6  
AAW14202

ID AAW14202 standard; protein: 199 AA.

XX

XX

AC

AC

25-JUN-1997 (first entry)

XX

XX

DE

Mature pokeweed antiviral protein deletion variant (1-199).

XX

XX

OS

Phytolacca americana.

XX

XX

PN

W09703183-Al.

XX

XX

PD

30-JAN-1997.

XX

XX

PF

11-JUL-1996; 96WO-US11546.

XX

XX

PP

11-JUL-1995; 95US-0500694.

XX

XX

PR

11-JUL-1995; 95US-0500611.

XX

XX

PA

(RUTF ) UNIV RUTGERS STATE NEW JERSEY.

XX

XX

PT

New pokeweed antiviral protein mutants - which exhibit anti-viral

and anti-fungal activity in plants and have reduced phytotoxicity

XX

XX

PS

Tumor NE;

XX

XX

WPI: 1997-119040/11.

XX

XX

CC

Proteins AAW14163-W14256 represent new pokeweed antiviral protein (PAP)

mutants having reduced phytotoxicity but retaining anti-viral or

anti-fungal activity in plants. The sequences' numbering corresponds to

the 262 amino acid mature protein. The proteins AAW14163-W14217

causing plant cell death or sterility. They can also confer increased resistance to insects, bacteria and nematodes in plants. N.B. This sequence is not given in the specification but is generated from the wild type sequence reproduced in the specification.

Sequence 196 AA:

Query Match 9.6%; Score 30; DB 18; Length 196;  
Best Local Similarity 100.0%; Pred. No. 1.9e-20;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 190 FLLVAIQMVSEAAAFKYEINOVKTNFNAF 219  
Db 166 FLLVAIQMVSEAAAFKYEINOVKTNFNAF 195

# RESULT 4

AAW14204  
D AAW14204 standard; protein: 197 AA.

XX AAW14204;

25-JUN-1997 (first entry)

Mature pokeweed antiviral protein deletion variant (1-197).

Pokeweed antiviral protein; mutant; phytotoxicity; anti-viral; virus; anti-fungal; fungus; plant; ribosomal inhibitor protein; adenine; insect; stem-loop; 28S rRNA; eukaryotic; ribosome; elongation factor-2 binding; translation; broad spectrum; resistance; cell death; sterility; nematode; bacterium.

Phytolacca americana.

W09703183-A1.

30-JAN-1997.

11-JUL-1996; 96WO-US11546.

11-JUL-1995; 95US-0500694.

11-JUL-1995; 95US-0500611.

(RUTE ) UNIV RUTGERS STATE NEW JERSEY.

Tuner NE;

WPI: 1997-119040/11.

New pokeweed antiviral protein mutants - which exhibit anti-viral and anti-fungal activity in plants and have reduced phytotoxicity

Claim 8: -; 64pp; English.

Proteins AAW14163-W14256 represent new pokeweed antiviral protein (PAP) mutants having reduced phytotoxicity but retaining anti-viral or anti-fungal activity in plants. The sequences' numbering corresponds to the 262 amino acid mature protein. The proteins AAW14165-W14217 represent serial deletions from the C-terminus whereas sequences AAW14218-55 are serial deletions from the N-terminus. Prior to processing, the protein contains a 22 amino acid N-terminal signal sequence and a 29 amino acid C-terminal extension. PAP is a Type I ribosomal inhibitor protein (RIP) that catalyses the removal of a specific adenine residue from a highly conserved stem-loop structure in the 28S rRNA of eukaryotic ribosomes and also interferes with elongation factor-2 binding thus blocking protein translation. The PAP mutants can confer broad spectrum virus and fungus resistance to plants without causing plant cell death or sterility. They can also confer increased resistance to insects, bacteria and nematodes in plants. N.B. This sequence is not given in the specification but is generated from the wild type sequence reproduced in the specification.

Sequence. 197 AA:

Query Match 9.6%; Score 30; DB 18; Length 198;  
Best Local Similarity 100.0%; Pred. No. 1.9e-20;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 190 FLLVAIQMVSEAAAFKYEINOVKTNFNAF 219

Query Match 9.6%; Score 30; DB 18; Length 197;  
Best Local Similarity 100.0%; Pred. No. 1.9e-20;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 190 FLLVAIQMVSEAAAFKYEINOVKTNFNAF 219  
Db 166 FLLVAIQMVSEAAAFKYEINOVKTNFNAF 195

# RESULT 5

AAW14203  
D AAW14203 standard; protein: 198 AA.

XX AAW14203;

25-JUN-1997 (first entry)

Mature pokeweed antiviral protein deletion variant (1-198).

Pokeweed antiviral protein; mutant; phytotoxicity; anti-viral; virus; anti-fungal; fungus; plant; ribosomal inhibitor protein; adenine; insect; stem-loop; 28S rRNA; eukaryotic; ribosome; elongation factor-2 binding; translation; broad spectrum; resistance; cell death; sterility; nematode; bacterium.

Phytolacca americana.

W09703183-A1.

30-JAN-1997.

11-JUL-1996; 96WO-US11546.

11-JUL-1995; 95US-0500694.

11-JUL-1995; 95US-0500611.

(RUTE ) UNIV RUTGERS STATE NEW JERSEY.

Tuner NE;

WPI: 1997-119040/11.

New pokeweed antiviral protein mutants - which exhibit anti-viral and anti-fungal activity in plants and have reduced phytotoxicity

Claim 8: -; 64pp; English.

Proteins AAW14163-W14256 represent new pokeweed antiviral protein (PAP) mutants having reduced phytotoxicity but retaining anti-viral or anti-fungal activity in plants. The sequences' numbering corresponds to the 262 amino acid mature protein. The proteins AAW14165-W14217 represent serial deletions from the C-terminus whereas sequences AAW14218-55 are serial deletions from the N-terminus. Prior to processing, the protein contains a 22 amino acid N-terminal signal sequence and a 29 amino acid C-terminal extension. PAP is a Type I ribosomal inhibitor protein (RIP) that catalyses the removal of a specific adenine residue from a highly conserved stem-loop structure in the 28S rRNA of eukaryotic ribosomes and also interferes with elongation factor-2 binding thus blocking protein translation. The PAP mutants can confer broad spectrum virus and fungus resistance to plants without causing plant cell death or sterility. They can also confer increased resistance to insects, bacteria and nematodes in plants. N.B. This sequence is not given in the specification but is generated from the wild type sequence reproduced in the specification.

Sequence 198 AA:

New pokeweed antiviral protein (PAP) with similar activity to ricin - used to treat cancer and as an agricultural chemical

Claim 1: Page 11-13; 14pp; Japanese.

NB: A protein comprising 261 amino acids is claimed. PAP has a similar activity to ricin, i.e. inhibits protein synthesis. The protein may be obtained all year round by recombinant DNA techniques. PAP can be used partic. against cancer and as an agricultural chemical.

Total mRNA is extracted from the seeds, leaves and roots of pokeweed and used to prepare cDNA using PCR. The resultant cDNA is used to prepare two DNA fractions, which are introduced into a cloning vector EMBL3 and then into host E.coli PLK-17 (P2) to produce PAP.

Query Match 10.2%; Score 32; DB 14; Length 294;  
Best Local Similarity 100.0%; Pred. No. 3.4e-22;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

190 FLVAIQMVSEAAAREKYIENQVKTENRAFYP 221  
189 FLVAIQMVSEAAAREKYIENQVKTENRAFYP 220

RESULT 2  
AAW14206  
ID AAW14206 standard; protein: 195 AA.

XX AAW14206;

25-JUN-1997 (first entry)

Mature pokeweed antiviral protein deletion variant (1-195).

Pokeweed antiviral protein; mutant; phytoxicity; anti-viral; virus; anti-fungal; fungus; plant; ribosomal inhibitor protein; adenine; insect; stem-loop; 28S rRNA; eukaryotic; ribosome; elongation factor-2 binding; translation; broad spectrum; resistance; cell death; sterility; nematode; bacterium.

Phytolacca americana.

WO9703183-A1.

30-JAN-1997.

11-JUL-1996; 96WO-US11546.

11-JUL-1995; 95US-0500694.

11-JUL-1995; 95US-0500611.

(RUTF) UNIV RUTGERS STATE NEW JERSEY.

Tuner NE;

WPI: 1997-119040/11.

New pokeweed antiviral protein mutants - which exhibit anti-viral and anti-fungal activity in plants and have reduced phytoxicity

Claim 8: -; 64pp; English.

Proteins AAW14163-W14256 represent new pokeweed antiviral protein (PAP) mutants having reduced phytoxicity but retaining anti-viral or anti-fungal activity in plants. The sequences, numbering corresponds to the 262 amino acid mature protein. The proteins AAW14165-W14217 represent serial deletions from the C-terminus whereas sequences AAW14218-55 are serial deletions from the N-terminus. Prior to processing, the protein contains a 22 amino acid N-terminal signal

sequence and a 29 amino acid C-terminal extension. PAP is a Type I ribosomal inhibitor protein (RIP) that catalyses the removal of a specific adenine residue from a highly conserved stem-loop structure in the 28S rRNA of eukaryotic ribosomes and also interferes with elongation factor-2 binding thus blocking protein translation. The PAP mutants can confer broad spectrum virus and fungus resistance to plants without causing plant cell death or sterility. They can also confer increased resistance to insects, bacteria and nematodes in plants. N.B. This sequence is not given in the specification but is generated from the wild type sequence reproduced in the specification.

Sequence 195 AA;

Query Match 9.6%; Score 30; DB 18; Length 195;  
Best Local Similarity 100.0%; Pred. No. 1.9e-20;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

190 FLVAIQMVSEAAAREKYIENQVKTENRAF 219  
166 FLVAIQMVSEAAAREKYIENQVKTENRAF 195

RESULT 3

AAW14205

ID AAW14205 standard; protein: 196 AA.

XX AAW14205;

25-JUN-1997 (first entry)

Mature pokeweed antiviral protein deletion variant (1-196).

Pokeweed antiviral protein; mutant; phytoxicity; anti-viral; virus; anti-fungal; fungus; plant; ribosomal inhibitor protein; adenine; insect; stem-loop; 28S rRNA; eukaryotic; ribosome; elongation factor-2 binding; translation; broad spectrum; resistance; cell death; sterility; nematode; bacterium.

Phytolacca americana.

WO9703183-A1.

30-JAN-1997.

11-JUL-1996; 96WO-US11546.

11-JUL-1995; 95US-0500694.

11-JUL-1995; 95US-0500611.

(RUTF) UNIV RUTGERS STATE NEW JERSEY.

Tuner NE;

WPI: 1997-119040/11.

New pokeweed antiviral protein mutants - which exhibit anti-viral and anti-fungal activity in plants and have reduced phytoxicity

Claim 8: -; 64pp; English.

Proteins AAW14163-W14256 represent new pokeweed antiviral protein (PAP) mutants having reduced phytoxicity but retaining anti-viral or anti-fungal activity in plants. The sequences, numbering corresponds to the 262 amino acid mature protein. The proteins AAW14165-W14217 represent serial deletions from the C-terminus whereas sequences AAW14218-55 are serial deletions from the N-terminus. Prior to processing, the protein contains a 22 amino acid N-terminal signal sequence and a 29 amino acid C-terminal extension. PAP is a Type I ribosomal inhibitor protein (RIP) that catalyses the removal of a specific adenine residue from a highly conserved stem-loop structure in the 28S rRNA of eukaryotic ribosomes and also interferes with elongation factor-2 binding thus blocking protein translation. The PAP mutants can confer broad spectrum virus and fungus resistance to plants without

GenCore version 5.1.4\_p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

protein - protein search, using sw model

un on: May 28, 2003, 09:59:10 : Search time 77 seconds  
(without alignments)  
543,386 Million cell updates/sec

Title: US-09-978-274A-2  
Perfect score: 314  
Sequence: 1 MKVMLVVVTLIAWLAAPT.....VIISTYYNMSNLGLDFEGF 314

coring table: OLIGO  
Gapop 60.0 , Capext 60.0

searched: 908470 seqs, 133250620 residues

ord size : 0

total number of hits satisfying chosen parameters: 908470

inimum DB seq length: 0  
aximum DB seq length: 2000000000

ost-processing: Listing first 45 summaries

atabase : A\_Geneseq\_101002.\*

- 1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*
- 2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*
- 3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*
- 4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*
- 5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*
- 6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*
- 7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*
- 9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*
- 10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*
- 11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*
- 12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*
- 13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*
- 14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*
- 15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*
- 16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*
- 17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*
- 18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*
- 19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*
- 20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*
- 21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*
- 22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*
- 23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

ID	Query Match	Length	DB	Description
1	32	10-2	294	14 AAR37345 Pokeweed antiviral
2	30	9-6	195	18 AAW14206 Mature pokeweed an
3	30	9-6	196	18 AAW14205 Mature pokeweed an
4	30	9-6	197	18 AAW14204 Mature pokeweed an
5	30	9-6	198	18 AAW14203 Mature pokeweed an
6	30	9-6	199	18 AAW14202 Mature pokeweed an
7	30	9-6	200	18 AAW14201 Mature pokeweed an
8	30	9-6	201	18 AAW14200 Mature pokeweed an
9	30	9-6	202	18 AAW14199 Mature pokeweed an
10	30	9-6	203	18 AAW14198 Mature pokeweed an

11	30	9-6	204	18 AAW14197 Mature pokeweed an
12	30	9-6	205	18 AAW14196 Mature pokeweed an
13	30	9-6	206	18 AAW14195 Mature pokeweed an
14	30	9-6	207	18 AAW14194 Mature pokeweed an
15	30	9-6	208	18 AAW14193 Mature pokeweed an
16	30	9-6	209	18 AAW14192 Mature pokeweed an
17	30	9-6	210	18 AAW14191 Mature pokeweed an
18	30	9-6	211	18 AAW14190 Mature pokeweed an
19	30	9-6	212	18 AAW14189 Mature pokeweed an
20	30	9-6	213	18 AAW14188 Mature pokeweed an
21	30	9-6	214	18 AAW14187 Mature pokeweed an
22	30	9-6	215	18 AAW14186 Mature pokeweed an
23	30	9-6	216	18 AAW14185 Mature pokeweed an
24	30	9-6	217	18 AAW14184 Mature pokeweed an
25	30	9-6	218	18 AAW14183 Mature pokeweed an
26	30	9-6	219	18 AAW14182 Mature pokeweed an
27	30	9-6	220	18 AAW14181 Mature pokeweed an
28	30	9-6	221	18 AAW14180 Mature pokeweed an
29	30	9-6	222	18 AAW14179 Mature pokeweed an
30	30	9-6	223	18 AAW14178 Mature pokeweed an
31	30	9-6	224	18 AAW14177 Mature pokeweed an
32	30	9-6	225	18 AAW14176 Mature pokeweed an
33	30	9-6	226	18 AAW14175 Mature pokeweed an
34	30	9-6	227	18 AAW14174 Mature pokeweed an
35	30	9-6	228	18 AAW14173 Mature pokeweed an
36	30	9-6	229	18 AAW14172 Mature pokeweed an
37	30	9-6	230	18 AAW14171 Mature pokeweed an
38	30	9-6	231	18 AAW14170 Mature pokeweed an
39	30	9-6	232	18 AAW14169 Mature pokeweed an
40	30	9-6	233	18 AAW14168 Mature pokeweed an
41	30	9-6	234	18 AAW14167 Mature pokeweed an
42	30	9-6	235	18 AAW14166 Mature pokeweed an
43	30	9-6	236	18 AAW14165 Mature pokeweed an
44	30	9-6	237	18 AAW14164 Mature pokeweed an
45	30	9-6	238	18 AAW14163 Mature pokeweed an

ALIGNMENTS

RESULT 1  
AAR37345  
ID AAR37345 standard; Protein: 294 AA.  
XX AC AAR37345;  
XX DT 09-NOV-1993 (first entry)  
XX DE Pokeweed antiviral protein.  
XX KW Pokeweed: ricin; protein synthesis inhibitor; cancer;  
XX KW polymerase chain reaction; PCR.  
XX OS Phytolacca americana.  
XX FT Key Location/Qualifiers  
XX FT Peptide 1..24  
XX FT Protein /label= sig\_peptide  
XX FT /label= mat\_protein  
XX PN JP05137580-A.  
XX PD 01-JUN-1993.  
XX PF 20-NOV-1991; 91JP-0325672.  
XX PR 20-NOV-1991; 91JP-0325672.  
XX PA (NIDB ) JAPAN TOBACCO INC.  
XX DR WPI: 1593-211306/26.  
XX DR N-PSDB: AAO43967.



OY 1 MKVMLVVVVTLIAHLIAAPTSTCAINTITDAGNATINKYATFMSLRNOAKDPKLCY 60  
DB 1 MKVMLVVVVTLIAHLIAAPTSTCAINTITDAGNATINKYATFMSLRNOAKDPKLCY 60  
OY 61 IPMLPDTNSTPKYLLVLOGANLKTITMLRRNNLYVMGYSDPFNGNCKRYHIFNDITST 120  
DB 61 IPMLPDTNSTPKYLLVLOGANLKTITMLRRNNLYVMGYSDPFNGNCKRYHIFNDITST 120  
OY 121 ERTDVENTLSSSSSRVAMSINYSLYPTMEKKAENSRNOVOLGIGILSSDICKISGVD 180  
DB 121 ERTDVENTLSSSSSRVAMSINYSLYPTMEKKAENSRNOVOLGIGILSSDICKISGVD 180  
OY 181 SPVPKTEAFLLVAIQWSEAAARFYIENQVKTNFNRAFPDPKVINLEEKWKISIAH 240  
DB 181 SPVPKTEAFLLVAIQWSEAAARFYIENQVKTNFNRAFPDPKVINLEEKWKISIAH 240  
OY 241 NAKNGALPKPLELVDAKGTWIVLRVDEINRDVALLKYVNGTCOTTYONAMFQVISTY 300  
DB 241 NAKNGALPKPLELVDAKGTWIVLRVDEINRDVALLKYVNGTCOTTYONAMFQVISTY 300  
OY 301 YNYMSNLGDLFEFG 314  
DB 301 YNYMSNLGDLFEFG 314

RESULT 2  
Q8S946 PRELIMINARY: PRT: 262 AA.  
AC Q8S946:  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DE PAP-S2 (fragment).  
CN PAPS2.  
OS *Phytolacca americana* (Common pokeweed) (Virginalian pokeweed)  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Caryophyllidae; Caryophyllales; Phytolaccaceae; Phytolacca.  
OX NCBI\_Taxid=3527;  
RN 111  
KP Phytolacca americana (Common pokeweed)  
RA Honjo E., Watanabe K.;  
RT "Cloning of genomic DNA encoding two types of pokeweed antiviral  
RT protein in seeds, PAP-S1 and PAP-S2, and functional comparison of  
RT their recombinant proteins with other PAP isoforms."  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB071855; BAB86350.1;  
FT NON\_TER 1 1  
FT NON\_TER 262 262  
SQ SEQUENCE 262 AA; 29486 MW; AF2D010A73C9D18B CRC64;

Query Match 82.4%; Score 1332; DB 10; Length 262;  
Best Local Similarity 98.1%; Pred. No. 2.4e-96;  
Matches 257; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
OY 25 INTITDAGNATINKYATFMSLRNOAKDPKLCYGIPLPOTNSTPKYLLVLOGANLK 84  
DB 1 INTITDAGNATINKYATFMSLRNOAKDPKLCYGIPLPOTNSTPKYLLVLOGANLK 60  
OY 85 TITLMLRRNNLYVMGYSDPFNGNCKRYHIFNDITSTERTDVENTLSSSSSRVAMSINYN 144  
DB 61 TITLMLRRNNLYVMGYSDPFNGNCKRYHIFNDITSTERTDVENTLSSSSSRVAMSINYN 120  
OY 145 SLIPTMEKKAENSRNOVOLGIGILSSDICKISGVDSPVPKTEAFLLVAIQWSEAAARF 204  
DB 121 SLIPTMEKKAENSRNOVOLGIGILSSDICKISGVDSPVPKTEAFLLVAIQWSEAAARF 180  
OY 205 KYIENQVKTNFNRAFPDPKVINLEEKWKISIAHNAKNGALPKPLELVDAKGTWIVL 264  
DB 181 KYIENQVKTNFNRAFPDPKVINLEEKWKISIAHNAKNGALPKPLELVDAKGTWIVL 240  
OY 265 RVDEINRDVALLKYVNGTCQT 286

DB 241 RVDEINRDVALLKYVNGTCQT 262  
RESULT 3  
Q941G8 PRELIMINARY: PRT: 313 AA.  
AC Q941G8:  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE rRNA -glycosidase (EC 3.2.2.22).  
GN PAP.  
OS *Phytolacca acinosa*.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Caryophyllidae; Caryophyllales; Phytolaccaceae; Phytolacca.  
OX NCBI\_Taxid=107615;  
RN 111  
RP SEQUENCE FROM N.A.  
RA Peng X., Yuan J., Qiang B.;  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
CC -!- SPECIFIC ADENOSINE ON THE 28S rRNA.  
CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY.  
DR EMBL; AY045785; AAL15442.1;  
DR InterPro: IPR001574; R1P.  
DR Pfam: PF00161; R1P; 1.  
KW Hydrolase; Toxin.  
SQ SEQUENCE 313 AA; 35059 MW; E478ED571C17885 CRC64;

Query Match 76.2%; Score 1231.5; DB 10; Length 313;  
Best Local Similarity 75.9%; Pred. No. 2.2e-88;  
Matches 239; Conservative 33; Mismatches 40; Indels 3; Gaps 2;  
OY 1 MKVMLVVVVTLIAHLIAAPTSTCAINTITDAGNATINKYATFMSLRNOAKDPKLCY 60  
DB 1 MKSMC--VVTISVMCLAPTSTWAVNTIYNGVSTTISKYATFLDLNRNEAKDPSLKCY 58  
OY 61 IPMLPDTNSTPKYLLVLOGANLKTITMLRRNNLYVMGYSDPFNGNCKRYHIFNDITST 120  
DB 59 IPMLPDTNSTPKYLLVLOGANLKTITMLRRNNLYVMGYSDPFNGNCKRYHIFNDITST 118  
OY 121 ERTDVENTLSSSSSRVAMSINYSLYPTMEKKAENSRNOVOLGIGILSSDICKISGVD 180  
DB 119 ERQDVTEFLTCPNPSRVSKNINYSRPTLESKAGVSKSVOVLGIGILSSDICKISGVT 178  
OY 181 SPVPKTEAFLLVAIQWSEAAARFYIENQVKTNFNRAFPDPKVINLEEKWKISIAH 240  
DB 179 SFTEKTEAFLLVAIQWSEAAARFYIENQVKTNFNRAFPDPKVINLEEKWKISIAH 238  
OY 241 NAKNGALPKPLELVDAKGTWIVLRVDEINRDVALLKYVNGTCOTTYONAMFQVISTY 299  
DB 239 DAKNGVLPKPLELVDAKGTWIVLRVDDIKPUVALNLYVGGSCOTTYNONAMFPOLIMST 298  
OY 300 YNYMSNLGDLFEFG 314  
DB 299 YNYMANLGDLPFEG 313

RESULT 4  
Q9XFF8 PRELIMINARY: PRT: 315 AA.  
AC Q9XFF8:  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE rRNA -glycosidase (EC 3.2.2.22).  
GN PIP2.  
OS *Phytolacca insularis*.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Caryophyllidae; Caryophyllales; Phytolaccaceae; Phytolacca.

Query Match 82.4%; Score 1332; DB 10; Length 262;  
Best Local Similarity 98.1%; Pred. No. 2.4e-96;  
Matches 257; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
OY 25 INTITDAGNATINKYATFMSLRNOAKDPKLCYGIPLPOTNSTPKYLLVLOGANLK 84  
DB 1 INTITDAGNATINKYATFMSLRNOAKDPKLCYGIPLPOTNSTPKYLLVLOGANLK 60  
OY 85 TITLMLRRNNLYVMGYSDPFNGNCKRYHIFNDITSTERTDVENTLSSSSSRVAMSINYN 144  
DB 61 TITLMLRRNNLYVMGYSDPFNGNCKRYHIFNDITSTERTDVENTLSSSSSRVAMSINYN 120  
OY 145 SLIPTMEKKAENSRNOVOLGIGILSSDICKISGVDSPVPKTEAFLLVAIQWSEAAARF 204  
DB 121 SLIPTMEKKAENSRNOVOLGIGILSSDICKISGVDSPVPKTEAFLLVAIQWSEAAARF 180  
OY 205 KYIENQVKTNFNRAFPDPKVINLEEKWKISIAHNAKNGALPKPLELVDAKGTWIVL 264  
DB 181 KYIENQVKTNFNRAFPDPKVINLEEKWKISIAHNAKNGALPKPLELVDAKGTWIVL 240  
OY 265 RVDEINRDVALLKYVNGTCQT 286

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 28, 2003, 09:51:54 ; Search time 87 Seconds  
(without alignments)  
743.664 Million cell updates/sec

US-09-978-274A-2

Perfect score: 1617

Sequence: 1 MKVMLVVVVTLIAWLIAPT.....VIITSYVYVMSNLGLFEGF 314

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- SPREMBL\_21.\*
- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_plant.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1617	100.0	314	P93444	Phytolacca
2	1332	82.4	262	O8S946	Phytolacca
3	1231.5	76.2	313	O941C8	Phytolacca
4	1147.5	71.0	315	O9XFF8	Phytolacca
5	1127	69.7	339	O8RIA4	Phytolacca
6	1099.5	68.0	261	O8S947	Phytolacca
7	953	58.9	237	O9ATB3	Phytolacca
8	702	43.4	302	P93077	Clerodendru
9	455.5	28.2	289	O93261	mesembryant
10	406	25.1	272	O39418	beta vulgar
11	393	24.9	279	O9M5K6	clenopodium
12	363.5	22.5	279	O9AUE3	chenopodium
13	363.5	22.5	294	O93Y65	dianthus ch
14	357.5	22.1	294	O93Y64	dianthus ch
15	349.5	21.6	294	O93Y66	dianthus ch
16	321.5	19.9	258	O9S9E4	gelonium mu

17	308	19.0	305	10	O8W4U4	longchavill
18	298	18.4	319	10	O8WVY2	spinacia of
19	292	18.1	541	10	O41174	ricinus com
20	274.5	17.0	279	10	O92T25	amaranthus
21	248.5	15.4	251	10	O48859	amaranthus
22	248	15.3	289	10	O94KE4	trichosan
23	247	15.3	300	10	O04356	iris hollan
24	247	15.3	300	10	O04357	iris hollan
25	246	15.2	278	10	O00980	infia cylin
26	245.5	15.2	270	10	O9SAQ5	amaranthus
27	245	15.2	270	10	O41611	trichosan
28	242.5	15.0	284	10	O96322	amaranthus
29	242.5	15.0	580	10	O94BW3	cinnamomum
30	242	15.0	566	10	O04072	sambucus ni
31	241	14.9	247	10	O9LRE3	trichosan
32	241	14.9	289	10	O41216	trichosan
33	240.5	14.9	581	10	O94BW5	cinnamomum
34	238	14.7	580	10	O94BW4	cinnamomum
35	237.5	14.7	580	10	O94BW7	cinnamomum
36	237.5	14.7	563	10	O945S2	sambucus ni
37	234.5	14.5	565	10	O04071	sambucus ni
38	234	14.5	604	10	O94654	polyonatum
39	233	14.4	298	10	O04158	iris hollan
40	232.5	14.4	286	10	O41257	momordica c
41	232.5	14.4	547	10	O946E9	abrus preca
42	230.5	14.3	252	10	O38761	abrus preca
43	228	14.1	549	10	O9FV22	cinnamomum
44	224.5	13.9	251	10	O96236	abrus preca
45	223.5	13.8	251	10	O96237	abrus preca

ALIGNMENTS

RESULT 1

P93444  
ID P93444 PRELIMINARY: PRT: 314 AA.  
AC P93444;  
DT 01-MAY-1997 (TrEMBLrel. 03, Created)  
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE RNA\_glycosidase precursor (EC 3.2.2.22).  
GN PAP-S.  
OS Phytolacca americana (Common pokeberry) (Virginian pokeweed).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Caryophyllales; Caryophyllaceae; Phytolaccaceae; Phytolacca.  
OX NCBI\_TaxID=3527;  
KN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=SEED;  
RX MEDLINE=97263479; PubMed=9109394;  
RA Poyet J.L., Hoeveler A.;  
RT "cDNA cloning of the gene encoding the Antiviral Protein from the  
seeds of Phytolacca americana and its expression in E.coli.";  
RL FEBS Lett. 406:97-100(1997).  
CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
-!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY.  
CC EMBL: X98079; CAA66702.1;  
DR HSP: Q03464; IAPA.  
DR InterPro: IPR001574; RIP.  
DR Pfam: PF00161; RIP; 1.  
DR PRINTS: PR00356; SHIGARICIN.  
DR PROSITE: PS00275; SHIGA\_RICIN; 1.  
KW Hydrolase; Signal; Toxin.  
FT SIGNAL 1 24 POTENTIAL.  
FT CHAIN 25 314 POTENTIAL.  
SQ SEQUENCE 314 AA: 35323 MW: 48953CES7789FF9E CRC64;

Query Match 100.0%; Score 1617; DB 10; Length 314;  
Best Local Similarity 100.0%; Pred. No. 1.6e-118;  
Matches 314; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 SLYPTMEKKAENVSRNOVQLGTOILSSDTGKISGVDSFPVKTEAFFLLVATOMVSEAAARF 204  
DB 120 GLYPTLEKKAGVTSRNEVOLGTOILSSKIGKISGOGSFTKTERDFLLVATOMVSEAAARF 179  
QY 205 KYIENOVKTNENRATYPPDKVINLEEKWKCTSEATHNAKNGALPKPLELVDAKGTWIVL 264  
DB 180 KYIENOVKTNENRDFSPNDKVLDEENWKCTSEATHNSKNGALPKPLELVDAKGTWIVL 239  
QY 265 RVDEINROVALLKYVNGTCQTT 286  
DB 240 RVDEIKPOVGLLNYNGTCQAT 261

Search completed: May 28, 2003, 09:56:05  
Job time : 78 secs

09978274

SRN

Printed 11/17/2004

06/02/2003

Db 240 RVDEIKPDVCLLYVNGTQCOAT 261

RESULT 14  
AAW21708  
ID AAW21708 standard: Protein; 261 AA.  
XX  
AC AAW21708;  
DT 26-SEP-1997 (first entry)  
XX  
DE PAP-S RIP.  
XX  
KW pro-Ribosome Inactivating Protein; proRIP; peptide linker; cancer;  
KW inactivation; eukaryotic ribosome; alpha fragment; beta fragment;  
KW inhibitor; protein synthesis; N-glycosidase; glycosidic bond; liver;  
KW rat; ribosomal 28S RNA; cellular proliferation; HIV-infected T cell.  
XX  
OS *Phytolacca americana*.  
XX  
FH Key  
FT Region  
FT Location/Qualifiers  
FT 151..161  
FT /note= "position of possible insertion of internal  
FT peptide linker sequence"  
XX  
XX U55635384-A.  
XX  
PN 03-JUN-1997.  
XX  
XX 11-JUN-1990; 90US-0535636.  
XX  
XX 09-DEC-1992; 92US-0987927.  
XX  
XX 11-JUN-1990; 90US-0535636.  
XX  
XX 26-JAN-1995; 95US-0378761.  
XX  
XX (DOWC ) DOWELANCO.  
XX  
XX Hey TD, Morgan ABR, Walsh TA;  
XX  
XX WPI: 1997-309831/28.  
XX  
XX Inactive precursor of maize ribosome-inactivating protein - also  
XX chimeric ribosome-inactivating protein precursors containing  
XX internal linker sequences  
XX  
XX Claim 2: Column 125-126; 121pp; English.  
XX  
XX The sequences given in AAW21698-710 represent Ribosome Inactivating  
XX Proteins (RIP's), which may be used in the construction of the  
XX proRIP of the invention. The proRIP has a selectively removable,  
XX internal peptide linker. The precursor sequence is incapable of  
XX inactivating eukaryotic ribosomes, but can be converted by removal  
XX of the linker into a protein having alpha and beta fragments and being  
XX capable of inactivating eukaryotic ribosomes. RIPs are potent  
XX inhibitors of eukaryotic protein synthesis. They possess a highly  
XX specific N-glycosidase activity which cleaves the glycosidic bond of  
XX adenine 4324 of rat liver ribosomal 28S RNA. RIP's selectively inhibit  
XX cellular proliferation of cells, e.g. cancer cells and HIV-infected T  
XX cells. The inactive proRIP proteins make it possible to provide protein  
XX synthesis inhibitors with uses in practical and improved ways not before  
XX possible. The RIP can be used to make cytotoxic conjugates.  
XX  
XX Sequence 261 AA;  
SQ

Query Match 67.8%; Score 1096.5; DB 18; Length 261;  
Best Local Similarity 82.8%; Pred. No. 31e-98;  
Matches 217; Conservative 14; Mismatches 30; Indels 1; Caps 1;  
Qy 25 INITITDAGNATINKYATFMESLRNOAKDPKLCYCIPIMLPDTNSTPKYLLVKKLOCANLX 84  
Db 1 INITITDAGNATINKYATFMESLRNEAKDPSLCYCIPIMLPNTNSTIKYLLVKKLOGASLK 60  
Qy 85 TITLMLRRNLIYVNGYSDPPNGNKKRYHIIFNDITSTERTDVENTLCSSSSRVAMSNYN 144

Db 61 TITLMLRRNLIYVNGYSDPYD-NKCRYHIIFNDIKTEYSDVENTLCPSSNPRVAKIPJNYN 119  
Qy 145 SLYPTMEKKAENSRNQVOLGIGIILSSDQKISGVDSFPVKTEAPFLLVAIONVSAARX 204  
Db 120 GLYPTLEKKACVTSRNEVOLGIGIILSSDQKISGVDSFTEKTEAKFELVAIONVSAARX 179  
Qy 205 KYIENOVKTNFRAFYDPKPVINLEEKWKQISEATHNAKNGALPKPLVLDKATKALX 264  
Db 180 KYIENOVKTNFRDFSPNDKVLDEENWCKJSTAIHNSKNGALPKPLVLDKATKALX 219  
Qy 265 RVDEIKPDVCLLYVNGTQCOAT 286  
Db 240 RVDEIKPDVCLLYVNGTQCOAT 261

RESULT 35  
AAW37298  
ID AAW37298 standard: protein; 261 AA.  
XX  
AC AAW37298;  
DT 13-SEP-1993 (first entry)  
XX  
DE Plant type I RIP pokeweed antiviral protein.  
XX  
KW Type I ribosome-inactivating protein; ricin; gelonin; momordin;  
KW immunoconjugate; autoimmune disease; cell killing; toxin; PAP.  
XX  
XX WC9309130-A.  
XX  
XX 13-MAY-1993.  
XX  
XX 04-NOV-1992; 92WQ-US09487.  
XX  
XX 04-NOV-1991; 91US-0787567.  
XX  
XX 19-JUN-1992; 92US-0901707.  
XX  
XX (XOMA ) XOMA CORP.  
XX  
XX Bernhard SL, Better MD, Carroll SF, Lane JA, Lei SP;  
XX WPI: 1993-167617/20.  
XX  
XX Analogues of type I ribosome inactivating protein - useful as  
XX cytotoxic agents, immuno toxins for treating autoimmune diseases,  
XX cancer, graft versus host disease and selective cell killing in-vivo  
XX  
XX Disclosure: Page 99-100; 163pp; English.  
XX  
XX The invention covers analogues of Type I RIPs. Pokeweed antiviral  
XX protein is a Type I RIP and the analogues of the invention have a  
XX Cys available for intermolecular disulphide bonding at an amino acid  
XX position corresp. to a position not naturally available for bonding;  
XX the Cys residue is located in the C-terminal region of the analogue  
XX between a position corresp. to amino acid 251 and the C-terminus of  
XX ricin A chain. The analogues are pref. joined via a disulphide  
XX linkage to a molecule which specifically binds to a target cell, e.g.  
XX an antibody fragment.  
XX  
XX Sequence 261 AA;  
SQ

Query Match 67.4%; Score 1089.5; DB 14; Length 261;  
Best Local Similarity 82.4%; Pred. No. 1.5e-97;  
Matches 216; Conservative 14; Mismatches 31; Indels 1; Caps 1;  
Qy 25 INITITDAGNATINKYATFMESLRNOAKDPKLCYCIPIMLPDTNSTPKYLLVKKLOCANLX 84  
Db 1 INITITDAGNATINKYATFMESLRNEAKDPSLCYCIPIMLPNTNSTIKYLLVKKLOGASLK 60  
Qy 85 TITLMLRRNLIYVNGYSDPPNGNKKRYHIIFNDITSTERTDVENTLCSSSSRVAMSNYN 144  
Db 61 TITLMLRRNLIYVNGYSDPYD-NKCRYHIIFNDIKTEYSDVENTLCPSSNPRVAKIPJNYN 119

PR 09-DEC-1992: 9205-0988430.  
PR 04-NOV-1991: 9105-0787567.  
PA (XOMA ) XOMA CORP.  
P1 Bernhard SL, Better MD, Carroll SF, Lane JA, Lei SP;  
XX WPI; 1995-193480/25.  
XX  
XX  
XX Polynucleotide(s) encoding gelonin analogues - having a cysteine  
PT residue for intermolecular bonding for the prodn. of immuno-toxins(s)  
XX  
XX Disclosure: Figure 8; 66pp; English.  
XX  
XX Analogues of Type I RIP are defined as non-naturally occurring  
CC polypeptides that share the ribosome-inactivating activity of the  
CC natural protein but differ in AA sequence. Preferred analogues have  
CC a Cys available for disulfide bonding located at a posn. It is AA  
CC sequence from the posn. corresp. to posn. 251 in ricin A-chain R7A  
CC to the carboxy terminus of the analogue. (AAR74176 is the sequence of  
CC ricin A-chain R7A, which is a Type II RIP). The primary AA  
CC sequence of the Type I RIPs gelonin, BRIP, momordin II, Luffin  
CC (see Islam et al. Agricultural Biological Chem. 54(5) 1343-45 1991),  
CC Alptrichosanthin (see Chow et al., J. Biol. Chem., 265, 8670-74  
CC 1990), momordin I (see Ho et al., BBA, 1088, 311-14 1991),  
CC Mirabilis antiviral protein (see Habuka et al., J. Biol. Chem.,  
CC 264(12) 6629-37 1989), pokeweed antiviral protein isolated from  
CC seeds (see Kung et al., Agric. Biol. Chem., 54(12), 3301-18 1990)  
CC and saporin (see Bonatti et al., Eur. J. Biochem., 183, 465-70  
CC 1989) are individually aligned with the primary sequence of the  
CC ricin A-chain (see Halling et al., Nucleic Acids Res., 13,  
CC 8019-8033 1985) respectively in Figures 1-9. The AAs invariant  
CC among the ricin A-chain and the Type I RIPs are indicated in Ft.  
XX  
SQ Sequence 261 AA;

Query Match 67.8%; Score 1096.5; DB 16; Length 261;  
Best Local Similarity 82.8%; Pred. No. 3.1e-98;  
Matches 21; Conservative 14; Mismatches 30; Indels 1; Gaps 1;  
Oy 25 INTTFDAGNATINKYATFMESELRNQAQDKPKLCYGIPLMDTNTSTPKYLLVKGQANLK 84  
Db 1 INTTFDAGNATINKYATFMESELRNQAQDKPKLCYGIPLMDTNTSTPKYLLVKGQANLK 60  
Oy 85 TITLMLRRNLYVMGYSDFPFGNKKRYHIFNDITSTERTDVENTLCSSSSSRVAMSYN 144  
Db 61 TITLMLRRNLYVMGYSDFPD-NKCRYHIFNDIKGTESYDVENTLCPSSNPRVAKPINYN 119  
Oy 145 SLYPTMEKKAENVSRNQVQLGIQILSSDGIKISGVDSFPVKTEAFFLLVAIOMVSEARF 204  
Db 120 GLYPTLEKKAGVTSRNEVQLGIQILSSDGIKISGVDSFTEKIEAKFLVIAIOMVSEARF 179  
Oy 205 KYIENQVKTENRAFYDPKPKVINLEEKWKISSEAHNAKNGALPKPLELVIAOMVSEARF 264  
Db 180 KYIENQVKTENRAFYDPKPKVINLEEKWKISSEAHNAKNGALPKPLELVIAOMVSEARF 239  
Oy 265 RVDEINRQVALLKYNVNGTCQPT 286  
Db 240 RVDEIKPDVGLLYVNGTCQAT 261

RESULT 13  
AAW25145  
ID AAW25145 standard; Protein: 261 AA.  
XX  
XX  
XX AAW25145;  
XX  
XX 02-DEC-1997 (first entry)  
XX  
XX PAP-S (a ribosome inhibitory protein) inactive precursor.  
XX  
XX Matze; proRIP; ribosome inactivating protein; alpha: beta subunit;  
KW internal linker; Barley Translation Inhibitor; Trichosanthin.  
KW

KW Ricin A-chain: Abrin-A A-chain: Saporin; SLT-1; Luffin A; MAP;  
KW Ricinus communis agglutinin; Momordin; PAP-S; Luffin-B; Dianthin 30;  
KW therapeutic toxin; tumour cell targeted; protein synthesis inhibitor;  
KW post-translational modification; cancer; neoplasia; HIV; AIDS;  
KW human immunodeficiency virus; acquired immune deficiency syndrome.  
XX  
XX Synthetic.  
XX US5646026-A.  
XX  
XX 08-JUL-1997.  
XX  
XX 11-JUN-1990: 90US-0535036.  
XX  
XX 09-DEC-1992: 92US-0987927.  
PR 11-JUN-1990: 90US-0535036.  
PR 26-JAN-1995: 95US-0378761.  
PR 07-JUN-1995: 95US-0485286.  
XX  
XX (DOWC ) DOWELANCO.  
XX  
XX Hey TD, Morgan AER, Walsh TA;  
XX WPI; 1997-362934/33.  
XX  
XX DNA encoding pro-ribosome inactivating proteins - inactive  
PT precursors of ribosome inactivating proteins; can be expressed in  
PT eukaryotic cells without causing cell death  
XX  
XX Claim 4: Column 125-128; 186pp; English.  
XX  
XX AAW25145 shows a PAP-S (a ribosome inactivating protein, RIP) protein  
CC that contains a selectively removable internal peptide linker sequence  
CC separating the alpha and beta units. When separated the two units regain  
CC activity and are capable of inactivating eukaryotic ribosomes and hence  
CC preventing protein production. Many different ribosome inhibitory  
CC proteins (RIPs) may be produced with an internal linker including  
CC maize RIP, Trichosanthin, Ricin A-chain, Abrin-A A-chain and  
CC Saporin. The RIPs can be used in the construction of therapeutic  
CC toxins targeted to specific cells such as tumour cells via the  
CC attachment of a targeting polypeptide, e.g. a monoclonal antibody.  
CC A further use is in HIV therapy (see US4869903). There is interest  
CC in expressing RIP recombinantly in host eukaryotic cells, because of  
CC the capacity to provide correct post-translational processing. However,  
CC RIPs effectively inhibit protein synthesis in eukaryotic cells resulting  
CC in cell death. Since the inactive RIP proteins are not cytotoxic to  
CC eukaryotic cells, they can be recombinantly expressed in such cells and  
CC then converted to active RIP proteins.  
XX  
SQ Sequence 261 AA;

Query Match 67.8%; Score 1096.5; DB 18; Length 261;  
Best Local Similarity 82.8%; Pred. No. 3.1e-98;  
Matches 21; Conservative 14; Mismatches 30; Indels 1; Gaps 1;  
Oy 25 INTTFDAGNATINKYATFMESELRNQAQDKPKLCYGIPLMDTNTSTPKYLLVKGQANLK 84  
Db 1 INTTFDAGNATINKYATFMESELRNQAQDKPKLCYGIPLMDTNTSTPKYLLVKGQANLK 60  
Oy 85 TITLMLRRNLYVMGYSDFPFGNKKRYHIFNDITSTERTDVENTLCSSSSSRVAMSYN 144  
Db 61 TITLMLRRNLYVMGYSDFPD-NKCRYHIFNDIKGTESYDVENTLCPSSNPRVAKPINYN 119  
Oy 145 SLYPTMEKKAENVSRNQVQLGIQILSSDGIKISGVDSFPVKTEAFFLLVAIOMVSEARF 204  
Db 120 GLYPTLEKKAGVTSRNEVQLGIQILSSDGIKISGVDSFTEKIEAKFLVIAIOMVSEARF 179  
Oy 205 KYIENQVKTENRAFYDPKPKVINLEEKWKISSEAHNAKNGALPKPLELVIAOMVSEARF 264  
Db 180 KYIENQVKTENRAFYDPKPKVINLEEKWKISSEAHNAKNGALPKPLELVIAOMVSEARF 239  
Oy 265 RVDEINRQVALLKYNVNGTCQPT 286  
Db 240 RVDEIKPDVGLLYVNGTCQAT 261

09978274

XX NB: A protein comprising 261 amino acids is claimed.  
CC PAP has a similar activity to ricin, i.e. inhibits protein synthesis.  
CC The protein may be obtained all year round by recombinant DNA  
CC techniques. PAP can be used partic. against cancer and as an  
CC agricultural chemical.  
CC Total mRNA, is extracted from the seeds, leaves and roots of  
CC pokeweed and used to prepare cDNA using PCR. The resultant cDNA is  
CC used to prepare two DNA fractions, which are introduced into a  
CC cloning vector EMBL3 and then into host E.coli PLK-17 (P2) to  
CC produce PAP.  
XX Sequence 294 AA:  
Query Match 68.9%; Score 1113.5; DB 14; Length 294;  
Best Local Similarity 75.4%; Pred. No. 8.3e-100;  
Matches 221; Conservative 27; Mismatches 42; Indels 3; Gaps 3;  
Qy 1 MKMVLVVVTLIAHLIAAPTSTCAINTIFDAGNATINKYATFMESLRNQADPKLKYC 60  
Db 1 MKMVLVVVWMLSLKLPSPSTWAINITFDVGNATINKYATFMESLRNQADPKLKYC 60  
Qy 61 IPMLPDTNSTPKYLLVLOGANLKTITLMLRRNLYVMGYSDPENGKCRYHIENDIT-S 119  
Db 61 IPMLPDTNSTPKYLLVLOGANLKTITLMLRRNLYVMGYSDPENGKCRYHIENDIT-S 119  
Qy 120 TERTDVENTLCSSSSRVMSINSLYPTMEKKAENSRNOVOLGIOLSSDICKISGV 179  
Db 120 TERTDVENTLCSSSSRVMSINSLYPTMEKKAENSRNOVOLGIOLSSDICKISGV 179  
Qy 180 DSPVPKTEAFLLVAIONVSEAAFPKYIENOVKTNFRAPYDPDPKVINLEEKWKISAI 239  
Db 180 DSPVPKTEAFLLVAIONVSEAAFPKYIENOVKTNFRAPYDPDPKVINLEEKWKISAI 239  
Qy 240 HNANGALPKLELVDAKGTWIVLRVDEINRDVALLKYVNGTCOTTYYONAMF 292  
Db 240 HNANGALPKLELVDAKGTWIVLRVDEINRDVALLKYVNGTCOTTYYONAMF 292  
RESULT 11  
AAR13112  
ID AAR13112 standard; protein: 261 AA.  
XX AC AAR13112;  
XX DT 08-OCT-1991 (first entry)  
XX DE Anti-viral protein.  
XX KW Protein synthesis inhibitor; plant virus.  
XX OS Phytolacca americana.  
XX FT Key Location/Qualifiers  
FT Disulfide-bond 84...105  
FT Disulfide-bond 34...258  
XX JPO3145498-A.  
XX PN 20-JUN-1991.  
XX PD 27-OCT-1989; 89JP-0281373.  
XX PF 27-OCT-1989; 89JP-0281373.  
XX PR (SUMO) SUMITOMO CHEM IND KK.  
XX PA WPI; 1991-226467/31.  
XX DR Bioactive protein of phytolacca amepicana - used as protein  
XX PT synthesis inhibitor against infection caused by plant virus.  
XX PS Claim 1: Fig 1: 6 pp: Japanese.

XX The sequence codes for a bioactive protein which can be used to  
CC inhibit infection by plant viruses. It is obtd. from the plant  
CC seeds but high yields and purity can be obtd. by synthesising the  
CC gene encoding it and introducing it into E.coli, yeast, or animal  
CC or plant cells.  
XX Sequence 261 AA:  
Query Match 67.8%; Score 1096.5; DB 12; Length 261;  
Best Local Similarity 82.8%; Pred. No. 3.1e-98;  
Matches 217; Conservative 14; Mismatches 30; Indels 1; Gaps 1;  
Qy 25 INTITFDAGNATINKYATFMESLRNQADPKLKYCIGYIPMLPDTNSTPKYLLVLOGANLK 84  
Db 1 INTITFDAGNATINKYATFMESLRNQADPKLKYCIGYIPMLPDTNSTPKYLLVLOGANLK 84  
Qy 85 TITLMLRRNLYVMGYSDPENGKCRYHIENDITSTERTDVENTLCSSSSRVMSINYN 144  
Db 61 TITLMLRRNLYVMGYSDPYD-NKCRYHIENDITSTERTDVENTLCSSSSRVMSINYN 119  
Qy 145 SLYPTMEKKAENSRNOVOLGIOLSSDICKISGVSDPENGKCRYHIENDITSTERTDVENTLCSSSSRVMSINYN 204  
Db 120 GLYPTLEKKAGVTSRNEVOLGIOLSSDICKISGVSDPENGKCRYHIENDITSTERTDVENTLCSSSSRVMSINYN 179  
Qy 205 KYIENOVKTNFRAPYDPDPKVINLEEKWKISAIHNAKNGALPKLELVDAKGTWIVL 264  
Db 180 KYIENOVKTNFRAPYDPDPKVINLEEKWKISAIHNAKNGALPKLELVDAKGTWIVL 239  
Qy 265 RVDEINRDVALLKYVNGTCOTT 286  
Db 240 RVDEIKPDVGLLYVNGTCOAT 261  
RESULT 12  
AAR74184  
ID AAR74184 standard; protein: 261 AA.  
XX AC AAR74184;  
XX DT 01-JAN-1996 (first entry)  
XX DE Type I RIP pokeweed anti-viral protein (PAP) isolated from seeds.  
XX KW Ribosome inactivating protein; RIP; Type I; cytotoxin; immunotoxin.  
XX OS Pokeweed.  
XX FT Key Location/Qualifiers  
FT Misc-difference 16  
FT /label= Invariant residue  
FT /note= "In ricin A-chain and the Type I RIPS"  
FT Misc-difference 24  
FT /label= see above  
FT Misc-difference 72  
FT /label= see above  
FT Misc-difference 122  
FT /label= see above  
FT Misc-difference 144  
FT /label= see above  
FT Misc-difference 175...176  
FT /label= see above  
FT Misc-difference 178  
FT /label= see above  
FT Misc-difference 207  
FT /label= see above  
XX US5416202-A.  
XX PN 16-MAY-1995.  
XX PD 09-DEC-1992; 92US-0988430.  
XX PF 09-DEC-1992; 92US-0988430.  
XX PS

SRNT

Printed 11/17/2004

Query Match 71.0%; Score 1147.5; DB 19; Length 315;  
Best local Similarity 72.2%; Pred. NO. 4.5e-103;  
Matches 229; Conservative 36; Mismatches 47; Indels 5; Gaps 4;

QY 1 MKVNLVVVVVTLIAWLIAPASTGCAI--NTITFDAGNATINKYATFMESLRNOAKDPKLC 58  
DB 1 MKMNVLVMTITAWLIAPASTWAASNPITFEVGNATINKYATFMESLRNOAKDPNLC 60  
QY 59 YGIMPLDPTNSTPKYLLVLOGANLKTITMLRRNNLYVMGYSDPPFNGNKCRIYHFDIT 118  
DB 61 YGIMPLDPTNSTPKYLLVLOGANLKTITMLRRNNLYVMGYADTYN-NKCRYHIFKDIS 119  
QY 119 -STERTDVENTLCSSSSRVAMSINYSLYPTMEKKAEVNSRNOVQIGIQLSSDICKIS 177  
DB 120 NTTERNOVMTLCPNMSRVKNTSYDSSYPALCKKVG-RSRKVKVQIGIQLNSDIKIY 178  
QY 178 GVDSPVVKTEAFFLLVAIQMVSEARFKYIENOVKTNFRAFYDPKVINLFEKWKISE 237  
DB 179 GVDVNEKTEAEFLVVAIQMVPEATRFKYIENOVKTNFRAFYDNKVLNLEETWGLIST 238  
QY 238 AIHNAKNGALPKLELDVDAKGTWKIVLRVDEINRQVALLKYVNGTCOTTYQNAMFSQV 297  
DB 239 AIHDAKNGALTLPLEINEDGTWKIVLRVDEIKPDVGLLVYVDGTCOTTYQSDMFPQRI 298  
QY 298 STYYNYSNLCGLDFEGF 314  
DB 299 STYYNIVNLGQDFEGF 315

RESULT 9  
AAM53033  
ID AAM53033 standard; Protein: 315 AA.  
XX AC AAM53033;  
DT 19-MAR-2002 (first entry)  
XX DE Phytolacca insularis antiviral protein (PIP).  
XX KW Antiviral protein; PIP; cloning vector.  
XX OS Phytolacca insularis.  
XX FH Key Location/Qualifiers  
FT Peptide 1..24  
FT Protein /label= Signal\_peptide  
FT Protein /label= Mature\_Pip  
FT Region 196..208  
FT /note= "This region is highlighted in the specification,  
FT indicating that it has some significance"

XX KR98021298-A.  
XX PD 25-JUN-1998.  
XX PF 16-SEP-1996; 96KR-0040110.  
XX PR 16-SEP-1996; 96KR-0040110.  
XX PA (JINR-) JIN RO LTD.  
XX PI Moon YH, Choi JN, Yoon YC, Jin JH, Hong EJ, Lee JH, Chang JI;  
XX PI Park YC, Choi GH, Kim CH, Kim MG, Song SG, Lee JS, Choi YD;  
XX DR WPI: 1999-299928/25.  
XX DR N-PSDB: ABA96543.  
XX XX Cloning vector using gene of Phytolacca insularis antiviral protein  
XX PT (PIP) -  
XX PS Claim 1: Fig 1; 9pp; Korean.

CC The invention relates to cloning vectors comprising a gene encoding  
CC the Phytolacca insularis antiviral protein, PIP (ABA96543), or  
CC a fragment thereof. The present sequence represents PIP.  
XX SQ Sequence 315 AA:  
Query Match 71.0%; Score 1147.5; DB 20; Length 315;  
Best local Similarity 72.2%; Pred. NO. 4.5e-103;  
Matches 229; Conservative 36; Mismatches 47; Indels 5; Gaps 4;

QY 1 MKVNLVVVVVTLIAWLIAPASTGCAI--NTITFDAGNATINKYATFMESLRNOAKDPKLC 58  
DB 1 MKMNVLVMTITAWLIAPASTWAASNPITFEVGNATINKYATFMESLRNOAKDPNLC 60  
QY 59 YGIMPLDPTNSTPKYLLVLOGANLKTITMLRRNNLYVMGYSDPPFNGNKCRIYHFDIT 118  
DB 61 YGIMPLDPTNSTPKYLLVLOGANLKTITMLRRNNLYVMGYADTYN-NKCRYHIFKDIS 119  
QY 119 -STERTDVENTLCSSSSRVAMSINYSLYPTMEKKAEVNSRNOVQIGIQLSSDICKIS 177  
DB 120 NTTERNOVMTLCPNMSRVKNTSYDSSYPALCKKVG-RSRKVKVQIGIQLNSDIKIY 178  
QY 178 GVDSPVVKTEAFFLLVAIQMVSEARFKYIENOVKTNFRAFYDPKVINLFEKWKISE 237  
DB 179 GVDVNEKTEAEFLVVAIQMVPEATRFKYIENOVKTNFRAFYDNKVLNLEETWGLIST 238  
QY 238 AIHNAKNGALPKLELDVDAKGTWKIVLRVDEINRQVALLKYVNGTCOTTYQNAMFSQV 297  
DB 239 AIHDAKNGALTLPLEINEDGTWKIVLRVDEIKPDVGLLVYVDGTCOTTYQSDMFPQRI 298  
QY 298 STYYNYSNLCGLDFEGF 314  
DB 299 STYYNIVNLGQDFEGF 315

RESULT 10  
AAR37345  
ID AAR37345 standard; Protein: 294 AA.  
XX AC AAR37345;  
DT 09-NOV-1993 (first entry)  
XX DE Pokeweed antiviral protein.  
XX KW Pokeweed; ricin; protein synthesis inhibitor; cancer;  
XX KW polymerase chain reaction; PCR.  
XX OS Phytolacca americana.  
XX FH Key Location/Qualifiers  
FT Peptide 1..24  
FT Protein /label= sig\_peptide  
FT Protein /label= mat\_protein  
XX JPO5137580-A.  
XX PD 01-JUN-1993.  
XX PF 20-NOV-1991; 91JP-0329672.  
XX PR 20-NOV-1991; 91JP-0329672.  
XX PA (NISR) JAPAN TOBACCO INC.  
XX DR WPI: 1993-211306/26.  
XX DR N-PSDB: AQA43967.  
XX XX New pokeweed antiviral protein (PAP) with similar activity to  
XX ricin - used to treat cancer and as an agricultural chemical  
XX PS Claim 1: Page 11-13; 14pp; Japanese.





DT 22-FEB-1995 (first entry)  
 XX Pokeweed Antiviral Protein mutant PAP-9.  
 DE Pokeweed Antiviral Protein; PAP; ribosome inactivating protein;  
 XX RIP type I; immunotoxin; cell targeting; mutant.  
 KW Phytolacca americana.  
 OS  
 XX Key Location/Qualifiers  
 FH Misc-difference 196  
 FT /note- "wild-type Phe is replaced by Tyr"  
 FT Misc-difference 211  
 FT /note- "wild-type Lys is replaced by Arg"  
 XX  
 XX FR2699553-A.  
 XX  
 XX 24-JUN-1994.  
 PD  
 XX 09-DEC-1992: 92PR-0014821.  
 XX  
 XX 09-DEC-1992: 92PR-0014821.  
 XX  
 XX (INNO-) INNOTHERAPIE LAB.  
 PA  
 XX Dore J, Gras E, Wijdenes J.  
 PI  
 XX WPI: 1994-227165/28.  
 DR  
 XX Gene encoding mutant protein toxic to eukaryotic but not  
 XX prokaryotic cells esp. mutant pokeweed antiviral protein - are  
 XX useful in immunotoxin prodn., protection against viral infection  
 XX and for targeted cell destruction  
 XX  
 XX Claim 10: Fig 1: 25pp: French.  
 PS  
 XX Mutants of type I Ribosome Inactivating Proteins which are toxic to  
 XX eukaryotic, but not prokaryotic, cells are preferably derived from  
 XX the wild-type Pokeweed Antiviral Protein (PAP) (AAR57153).  
 CC Specifically, the mutants are PAP1 (with Gly replacing wild-type  
 CC Arg at position 68) and PAP9 (with Tyr replacing wild-type Phe and  
 CC Arg replacing wild-type Lys at positions 196 and 211, respectively).  
 XX  
 XX Sequence 292 AA:

Query Match 72.0%; Score 1164.5; DB 15; Length 292;  
 Best Local Similarity 76.3%; Pred. No. 8.9e-105;  
 Matches 222; Conservative 32; Mismatches 36; Indels 1; Gaps 1:

QY 25 INTTFDAGNATINKYATFMESLRNOAKDPKLCYGIPLMDPTNSTPKYLLVLOGANLK 84  
 DB 2 VNTIYVNGSTTSKYATFLNDLRNEAKDPSLCYGIPLMDPTNTNPKYLLVLOGSNKK 61  
 QY 85 TITLRLRNLYVMGYSDPFNGNKRVIIFNDITSTERTDVENTLCSSSSRVAMSYN 144  
 DB 62 TITLRLRNLYVMGYSDPFNETNKRVIIFNDISGTQDQVETTLCPNANSRVSKNIFD 121  
 QY 145 SLYPTMEKKAENSRNOVQIGIQLISSDGIKTSVDSPFKVTEAFFLLVAIQMVSAAKF 204  
 DB 122 SRYPTLESKAGVRSQVQLGIQILDNSNIGISGYMSFTEKTEAEFLVAIQMVSAAKF 181  
 QY 205 KYIENQVKTNFNRAFYDPKPVNLEEKWKIISATHNKNGALPKPLELVDAKGTWIVL 264  
 DB 182 KYIENQVKTNFNRAFNPNKPVNLEQETWGRISTALHDKNVGLPKPLELVDAKGTWIVL 241  
 QY 265 RVDEINRQVALLKYVNGTCQTTY-QNAMESQVITSTYNYMNSLGLDFEGF 314  
 DB 242 RVDEIKPDVALLNYVGGSCQTTYNQNAMEPQLIMSTYNYMNSLGLDFEGF 292

RESULT 6  
 AAR57153  
 ID AAR57153 standard; protein; 292 AA.

XX AAR57153;  
 XX 22-FEB-1995 (first entry)  
 DE Pokeweed Antiviral Protein mutant PAP-1.  
 XX Pokeweed Antiviral Protein; PAP; ribosome inactivating protein;  
 KW RIP type I; immunotoxin; cell targeting; mutant.  
 OS  
 XX Phytolacca americana.  
 FH  
 XX Key Location/Qualifiers  
 FT Misc-difference 68  
 FT /note- "wild-type Arg is replaced by Gly"  
 XX  
 XX FR2699553-A.  
 XX  
 XX 24-JUN-1994.  
 PD  
 XX 09-DEC-1992: 92PR-0014821.  
 XX  
 XX 09-DEC-1992: 92PR-0014821.  
 XX  
 XX (INNO-) INNOTHERAPIE LAB.  
 PA  
 XX Dore J, Gras E, Wijdenes J.  
 PI  
 XX WPI: 1994-227165/28.  
 DR  
 XX Gene encoding mutant protein toxic to eukaryotic but not  
 XX prokaryotic cells esp. mutant pokeweed antiviral protein - are  
 XX useful in immunotoxin prodn., protection against viral infection  
 XX and for targeted cell destruction  
 XX  
 XX Claim 10: Fig 1: 25pp: French.  
 PS  
 XX Mutants of type I Ribosome Inactivating Proteins which are toxic to  
 XX eukaryotic, but not prokaryotic, cells are preferably derived from  
 XX the wild-type Pokeweed Antiviral Protein (PAP) (AAR57153).  
 CC Specifically, the mutants are PAP1 (with Gly replacing wild-type  
 CC Arg at position 68) and PAP9 (with Tyr replacing wild-type Phe and  
 CC Arg replacing wild-type Lys at positions 196 and 211, respectively).  
 XX  
 XX Sequence 292 AA:

Query Match 72.0%; Score 1163.5; DB 15; Length 292;  
 Best Local Similarity 76.6%; Pred. No. 1.1e-104;  
 Matches 223; Conservative 30; Mismatches 37; Indels 1; Gaps 1:

QY 25 INTTFDAGNATINKYATFMESLRNOAKDPKLCYGIPLMDPTNSTPKYLLVLOGANLK 84  
 DB 2 VNTIYVNGSTTSKYATFLNDLRNEAKDPSLCYGIPLMDPTNTNPKYLLVLOGSNKK 61  
 QY 85 TITLRLRNLYVMGYSDPFNGNKRVIIFNDITSTERTDVENTLCSSSSRVAMSYN 144  
 DB 62 TITLRLRNLYVMGYSDPFNETNKRVIIFNDISGTQDQVETTLCPNANSRVSKNIFD 121  
 QY 145 SLYPTMEKKAENSRNOVQIGIQLISSDGIKTSVDSPFKVTEAFFLLVAIQMVSAAKF 204  
 DB 122 SRYPTLESKAGVRSQVQLGIQILDNSNIGISGYMSFTEKTEAEFLVAIQMVSAAKF 181  
 QY 205 KYIENQVKTNFNRAFYDPKPVNLEEKWKIISATHNKNGALPKPLELVDAKGTWIVL 264  
 DB 182 KYIENQVKTNFNRAFNPNKPVNLEQETWGRISTALHDKNVGLPKPLELVDAKGTWIVL 241  
 QY 265 RVDEINRQVALLKYVNGTCQTTY-QNAMESQVITSTYNYMNSLGLDFEGF 314  
 DB 242 RVDEIKPDVALLNYVGGSCQTTYNQNAMEPQLIMSTYNYMNSLGLDFEGF 292

RESULT 7  
 AAR57153

AA848548:

04-SEP-1994 (first entry)

Sequence of Phytolacca antiviral protein (PAP).

Antiviral protein; PAP; virus-resistance; transgenic plant.

Phytolacca americana L.

EP585554-A.

09-MAR-1994.

30-JUN-1993; 93EP-0110445.

16-AUG-1992; 92KR-0014895.

(JINR-) JIN RO LTD.

(SHIN-) SHINRO KK.

Choi K, Jeon H, Jeong H, Kim M, Lee K, Moon Y, Na B;

WPI: 1994-076002/10.

N-PSDB; AAQ56672.

Expression vector for phytolacca antiviral protein - used for producing transgenic virus-resistant plants and for producing the antiviral agent

Disclosure: Fig 1; 15pp; English.

To isolate PAP gene, total cellular mRNA was purified from leaves of Phytolacca americana L. obtd. in Korea. A cDNA library was constructed. The PAP gene was selected by immunoscreening employing anti-PAP antibody. A deletion mutant was prepd. from the isolated PAP gene, and the DNA sequence of the PAP genomewas determined.

Sequence 313 AA:

Query Match 76.08; Score 1229.5; DB 15; Length 313;  
Best Local Similarity 75.68; Pred. No. 4.6e-111;  
Matches 238; Conservative 33; Mismatches 41; Indels 3; Gaps 2;  
Qy 1 MKVLLVVVTVTLIAIAPTSTCAINTITFDAGNATINKYATFMESLRQAKDKPKCYG 60  
Db 1 MKSLVVVTISI--WLLAPTSTWAVNTIYVVGSTTSKYATFLNDRNEAKDPSLKYG 58  
Qy 61 IPMLPDTNSTPKYLLVKGANLKTITLMLRRNNLYVMGYSDPPFNKCRHYHFNIDTST 120  
Db 59 IPMLPNTNTNPKYLLVKGANLKTITLMLRRNNLYVMGYSDPPFNKCRHYHFNIDTST 118  
Qy 121 ERTDVENTLCSSSSRVAMSYNSLYPTMEKKAQVNSRNOVOLGQIOLSSDICKISVD 180  
Db 119 ERQDVTTLCPNANSRVSKNFDSRYPTLESKAGVRSQVQLGQIOLDSNIGKISGV 178  
Qy 181 SFPVKTEAFFLLVATOMYSEARFYKIENQVKTENRFAFYDPKVINLEEKWKGISEAH 240  
Db 179 SFTKTEAEFLVLAQMYSEARFYKIENQVKTENRFAFNPNKVLNLQETWGIKSTAH 238  
Qy 241 NAKNGALPKPLVDKAGTKWIVLRVDEINRDVALLKYVNGTCOTTY-ONAMFSOVIIST 299  
Db 239 DAKNGVLPKPLVDASGAKWIVLRVDEIKPDVALLNVVGGSCOTTYNQAMFPOLIMST 298  
Qy 300 YNNYNSNLGDLFEFG 314  
Db 299 YNNYNSNLGDLFEFG 313

RESULT 4  
AAR57152  
ID AAR57152 standard; protein: 292 AA.

XX

AC AAR57152:

22-FEB-1995 (first entry)

Wild-type Pokeweed Antiviral Protein (PAP).

Pokeweed antiviral protein; PAP; ribosome inactivating protein;

RIP type I; Immunotoxin; cell targeting.

Phytolacca americana.

FR2699553-A.

24-JUN-1994.

09-DEC-1992; 92FR-0014821.

09-DEC-1992; 92FR-0014821.

(INNO-) INNOTHERAPIE LAB.

Dore J, Gras E, Wijdenes J;

WPI: 1994-227165/28.

Gene encoding mutant protein toxic to eukaryotic but not prokaryotic cells esp. mutant pokeweed antiviral protein - are useful in immunotoxin prodn... protection against viral infection and for targeted cell destruction

Disclosure: Fig 1; 25pp; French.

Mutants of type I Ribosome Inactivating Proteins which are toxic to eukaryotic, but not prokaryotic, cells are preferably derived from the wild-type Pokeweed Antiviral Protein (PAP) (AAR57152). Specifically, the mutants are PAP1 (with Gly replacing wild-type Arg at position 68) and PAP9 (with Tyr replacing wild-type Phe and Arg replacing wild-type Lys at positions 196 and 211, respectively).

Sequence 292 AA:

Query Match 72.48; Score 1170.5; DB 15; Length 292;  
Best Local Similarity 77.08; Pred. No. 2.3e-105;  
Matches 224; Conservative 30; Mismatches 36; Indels 1; Gaps 1;  
Qy 25 INTITFDAGNATINKYATFMESLRQAKDKPKCYGIPMLPDTNSTPKYLLVKGANLX 84  
Db 2 VNTIYVVGSTTSKYATFLNDRNEAKDPSLKYGIPMLPNTNTNPKYLLVLOGSNKK 61  
Qy 85 TITLMLRRNNLYVMGYSDPPFNKCRHYHFNIDTSTERTDVENTLCSSSSRVAMSYN 144  
Db 62 TITLMLRRNNLYVMGYSDPPFNKCRHYHFNIDTSTERTDVENTLCPNANSRVSKNIFD 121  
Qy 145 SLYPTMEKKAQVNSRNOVOLGQIOLSSDICKISGVSDPPVKTEAFFLLVATOMYSEAK 204  
Db 122 SRYPTLESKAGVRSQVQLGQIOLDSNIGKISGVMSFTEKTEAEFLVATOMYSEAK 181  
Qy 205 KYIENQVKTENRFAFYDPKVINLEEKWKGISEAHNAKNGALPKPLVDKAGTKWIV 264  
Db 182 KYIENQVKTENRFAFNPNKVLNLQETWGIKSTAJHDAGKNGVLPKPLVDASGAKWIV 241  
Qy 265 KVDEINRDVALLKYVNGTCOTTY-ONAMFSOVIISTYNNYNSNLGDLFEFG 314  
Db 242 KVDEIKPDVALLNVVGGSCOTTYNQAMFPOLIMSTYNNYNSNLGDLFEFG 292

RESULT 5  
AAR57154  
ID AAR57154 standard; protein: 292 AA.

XX  
AC AAR57154;  
XX

PT viruses and/or fungi -  
XX Examples: Page 4-5: 43pp; English.  
XX This is the pokeweed antiviral protein (PAP) amino acid sequence. PAP is  
CC a type I ribosome-inhibiting protein (RIP) found in the cell walls of  
CC Phytolacca americana (pokeweed). It is a single polypeptide chain that  
CC catalytically removes a specific adenine residue from a highly conserved  
CC stem-loop structure in the 28S rRNA of eukaryotic ribosomes. The pokeweed  
CC antiviral protein II (PAP II) protein confers antiviral and/or antifungal  
CC activities to plants. A DNA molecule encoding a PAP II protein with an  
CC intact catalytic active site amino acid residue (E172) is useful for  
CC generating transgenic plants. PAP II DNA is useful for generating a  
CC protoplast or introducing the DNA directly into a plant part prior to  
CC regeneration. Cells and seeds which exhibit anti-viral and/or anti-fungal  
CC activity thus have increased resistance to viruses and/or fungi. Viruses  
CC include potato virus X (PVX), cucumber mosaic virus (CMV), and tomato  
CC yellow leaf curl virus, and fungi include Pythium, Phytophthora,  
CC Sclerotinia and Aspergillus. PAP II proteins also confer resistance to  
CC other plants pests including insects, bacteria and nematodes. PAP II DNA  
CC is also useful for identifying a PAP II protein having reduced  
CC cytotoxicity. PAP II has a reduced cytotoxicity compared to PAP, so  
CC unlike PAP transgenic plants which are stunted and sterile. PAP II  
CC transgenic plants have a normal and fertile phenotype.  
XX Sequence 313 AA:

Query Match 76.4% Score 1235.5; DR 21; Length 313;  
Best Local Similarity 75.9%; Pred. No. 1.2e-111;  
Matches 239; Conservative 33; Mismatches 40; Indels 3; Gaps 2;  
Oy 1 MKVMLVVVTIAMIAPSTCAINTITFDAGNATINKYATFMSLRNOAKDKPKCYG 60  
Db 1 MKSMLVVTISI--WLLIAPTSTWAVNTIYVGVSTTISKYATFLNDRNEAKDPSLKCYG 58  
Oy 61 IPMLPDTNSTPKYLLVKGLOGANIKTITLMLRNNLYVMGYSDDPPNGKCRHYHNDITST 120  
Db 59 IPMLPNTNTNPKYVLVELOGSNKKTITLMLRNNLYVMGYSDDPPETNCRHYHNDISGT 118  
Oy 121 ERTDVENTLCSSSSSRVAMSTNSLYPTMEKKAEVNSRNQVQIQLSSDICKISGV 180  
Db 119 ERQDVETTLCPNANSRVSKNINFSRYPTLESKAGVRSQVQLCQIQLDSNICKISGV 178  
Oy 181 SFPVKTEAFFLLVAIONVSEARFKYIENQVKTNFNRAFNPNPKVNLQETWKGISTAIH 240  
Db 179 SFTKTEAEFLVAIONVSEARFKYIENQVKTNFNRAFNPNPKVNLQETWKGISTAIH 238  
Oy 241 NAKNCALPKPLEVDAGKTKWIVLRVDEINRDVALLKYVNGTCQTTY-ONAMFSOVITST 299  
Db 239 DAKNGVLPKPLEVDASGAKWIVLRVDEIKPDVALLNYVGGSCQTTYONAMFPOLTINST 298  
Oy 300 YNYNSNLGDLPEGF 314  
Db 299 YNYNVNGLDLPEGF 313

RESULT 2  
ID AAB36500 standard; Protein: 313 AA.  
XX AAB36500:  
XX AAB36500:  
XX 06-MAR-2001 (first entry)  
XX Phytolacca americana pokeweed antiviral protein SEQ ID NO:2.  
XX Phytolacca americana; pokeweed; pokeweed; pokeweed antiviral protein; PAP;  
KW cancer; bioregulatory; fusion protein; immunoconjugate; mutant;  
KW cytostatic; anti-HIV; human immunodeficiency virus; AIDS; leukaemia;  
KW lymphoma; brain tumour; neuroblastoma; soft tissue sarcoma;  
KW osteosarcoma.  
XX

OS Phytolacca americana.  
XX US6146628-A.  
XX 14-NOV-2000.  
XX 11-JUL-1995: 95US-0501253.  
XX 11-JUL-1995: 95US-0501253.  
PR (MINU ) UNIV MINNESOTA & RUTGERS.  
PA (UYNE-) UNIV STATE NEW JERSEY.  
XX Uckun FM, Tumor NE;  
XX WPI: 2001-040422/05.  
DR N-PSDB: AAC87929.  
XX Immunoconjugates useful for treating cancer and acquired  
PT immunodeficiency syndrome, comprises mutant pokeweed anti-viral protein  
PT and a targeting moiety that binds a cell surface receptor -  
XX  
PS Example 6: Column 49-52: 32pp; English.  
XX The present invention describes a fusion protein or an immunoconjugate  
CC (1), comprising mutant pokeweed Anti-viral Protein (PAP) having an  
CC amino acid substitution at residue 75, 97 or 176 of native PAP and a  
CC targeting moiety that binds a cell surface receptor. (1) can have  
CC cytostatic and anti-HIV activities, and is an inhibitor of cellular  
CC RNA or protein synthesis. (1) is useful for treating AIDS and cancers  
CC including leukemia, lymphoma, a brain tumour, neuroblastoma,  
CC osteosarcoma, soft tissue sarcoma, breast, prostate, ovarian,  
CC testicular, melanoma, lung, or colon cancer. Immunoconjugates prepared  
CC using PAP mutants exhibit an improved therapeutic index over  
CC immunoconjugates containing either wild-type PAP or variant PAP.  
CC The present sequence represents the wild-type PAP, which is given in  
CC the exemplification of the present invention.  
XX Sequence 313 AA:

Query Match 76.4% Score 1232.5; DR 22; Length 313;  
Best Local Similarity 75.6%; Pred. No. 2.4e-111;  
Matches 238; Conservative 34; Mismatches 40; Indels 3; Gaps 2;  
Oy 1 MKVMLVVVTIAMIAPSTCAINTITFDAGNATINKYATFMSLRNOAKDKPKCYG 60  
Db 1 MKSMLVVTISI--WLLIAPTSTWAVNTIYVGVSTTISKYATFLNDRNEAKDPSLKCYG 58  
Oy 61 IPMLPDTNSTPKYLLVKGLOGANIKTITLMLRNNLYVMGYSDDPPNGKCRHYHNDITST 120  
Db 59 IPMLPNTNTNPKYVLVELOGSNKKTITLMLRNNLYVMGYSDDPPETNCRHYHNDISGT 118  
Oy 121 ERTDVENTLCSSSSSRVAMSTNSLYPTMEKKAEVNSRNQVQIQLSSDICKISGV 180  
Db 119 ERQDVETTLCPNANSRVSKNINFSRYPTLESKAGVRSQVQLCQIQLDSNICKISGV 178  
Oy 181 SFPVKTEAFFLLVAIONVSEARFKYIENQVKTNFNRAFNPNPKVNLQETWKGISTAIH 240  
Db 179 SFTKTEAEFLVAIONVSEARFKYIENQVKTNFNRAFNPNPKVNLQETWKGISTAIH 238  
Oy 241 NAKNCALPKPLEVDAGKTKWIVLRVDEINRDVALLKYVNGTCQTTY-ONAMFSOVITST 299  
Db 239 DAKNGVLPKPLEVDASGAKWIVLRVDEIKPDVALLNYVGGSCQTTYONAMFPOLTINST 298  
Oy 300 YNYNSNLGDLPEGF 314  
Db 299 YNYNVNGLDLPEGF 313

RESULT 3  
ID AAB48548 standard; Protein: 313 AA.  
XX AAB48548  
XX

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 28, 2003, 09:45:44 : Search time 76 Seconds  
(without alignments)  
550.535 Million cell updates/sec

Title: US-09-978-274A-2

Perfect score: 1617

Sequence: 1 MKVMLVVVVTIAWLAIAAPT.....VIISTYNYNMLGDLFEGF 314

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_101002.\*  
1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1235.5	76.4	313	21 AAY58025	Wild-type pokeweed
2	1232.5	76.2	313	22 AAB36500	Phytolacca americana
3	1229.5	76.0	313	15 AAR48548	Sequence of Phytol
4	1170.5	72.4	292	15 AAR57152	Wild-type Pokeweed
5	1164.5	72.0	292	15 AAR57154	Pokeweed Antiviral
6	1163.5	72.0	292	15 AAR57153	Pokeweed Antiviral
7	1160.5	71.8	292	15 AAR57153	Phytolacca insular
8	1147.5	71.0	315	19 AAW26774	Phytolacca insular
9	1147.5	71.0	315	20 AAW51033	Phytolacca insular
10	1113.5	68.9	294	14 AAR37345	Pokeweed antiviral

11	1096.5	67.8	261	12 AAR13112	Anti-viral protein
12	1096.5	67.8	261	16 AAR74184	Type I RIP pokewee
13	1096.5	67.8	261	18 AAR25145	PAP-S (a ribosome
14	1096.5	67.8	261	14 AAR21708	PAP-S RIP, Phytol
15	1089.5	67.4	261	14 AAR37298	Plant type I RIP P
16	1089.5	67.4	261	16 AAR63910	Type I ribosome-in
17	1047	64.7	261	18 AAR14218	Mature pokeweed an
18	1046	64.7	262	18 AAR14164	Mature pokeweed an
19	1043	64.5	262	18 AAR14256	Mature pokeweed an
20	1041	64.4	260	18 AAR14219	Mature pokeweed an
21	1041	64.4	262	18 AAR14163	Mature pokeweed an
22	1036	64.1	259	18 AAR14220	Mature pokeweed an
23	1033	63.9	257	18 AAR14222	Mature pokeweed an
24	1033	63.9	258	18 AAR14221	Mature pokeweed an
25	1031	63.8	305	15 AAR54839	Antiviral protein
26	1030	63.7	256	18 AAR14223	Mature pokeweed an
27	1029	63.6	254	18 AAR14225	Mature pokeweed an
28	1029	63.6	255	18 AAR14224	Mature pokeweed an
29	1025	63.4	305	16 AAR80106	P. insularis anti
30	1023	63.3	253	18 AAR14226	Mature pokeweed an
31	1022	63.2	251	18 AAR14228	Mature pokeweed an
32	1022	63.2	252	18 AAR14227	Mature pokeweed an
33	1017	62.9	250	18 AAR14229	Mature pokeweed an
34	1013	62.6	249	18 AAR14230	Mature pokeweed an
35	1012	62.6	248	18 AAR14231	Mature pokeweed an
36	1007	62.3	247	18 AAR14232	Mature pokeweed an
37	1000	61.8	246	18 AAR14233	Mature pokeweed an
38	996	61.6	245	18 AAR14234	Mature pokeweed an
39	991	61.3	244	18 AAR14235	Mature pokeweed an
40	985	60.9	243	18 AAR14236	Mature pokeweed an
41	983	60.8	240	18 AAR14239	Mature pokeweed an
42	983	60.8	241	18 AAR14238	Mature pokeweed an
43	983	60.8	242	18 AAR14237	Mature pokeweed an
44	979	60.5	239	18 AAR14240	Mature pokeweed an
45	974	60.2	238	18 AAR14241	Mature pokeweed an

ALIGNMENTS

RESULT 1

AAY58025

ID AAY58025 standard; Protein: 313 AA.

XX AAY58025;

XX 29-FEB-2000 (first entry)

XX Wild-type pokeweed antiviral protein (PAP) amino acid sequence.

XX pokeweed antiviral protein; PAP; PAP II; antifungal; transgenic plant;

XX ribosome inhibiting protein; RIP; cereal crop; viral resistance; PVX;

XX potato virus X; cucumber mosaic virus; CMV;

XX tomato yellow leaf curl virus.

XX Phytolacca americana.

OS Phytolacca americana.

PN WO9960843-A1.

XX 02-DEC-1999.

XX 21-MAY-1999; 99MO-US11301.

XX 22-MAY-1998; 98US-U086374.

XX (RUTF ) UNIV KUTTERS STATE NEW JERSEY.

XX Turner NE, Wang P;

XX WPI; 2000-062555/05.

XX N-PSDR; AAZ45197.

XX New antiviral DNA useful for generating transgenic plants resistant to

Query Match 2.38, Score 22; DB 1; Length 918;  
Best Local Similarity 100.0%; Pred. No. 0.1;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 289 GTGATGGCTATTCTGATCCCT 310  
DB 277 GTGATGGCTATTCTGATCCCT 298

## RESULT 14

US-08-996-441B-51/C  
; Sequence 51, Application US/08996441B  
; Patent No. 6023013

## GENERAL INFORMATION:

APPLICANT: English, Leigh H.  
APPLICANT: Brussock, Susan M.  
APPLICANT: Malvar, Thomas M.  
APPLICANT: Bryson, James W.  
APPLICANT: Kulesza, Caroline A.  
APPLICANT: Walters, Frederick S.  
APPLICANT: Slatin, Stephen L.  
APPLICANT: Von Tersch, Michael A.  
APPLICANT: Romano, Charles

TITLE OF INVENTION: INSECT-RESISTANT TRANSGENIC PLANTS

NUMBER OF SEQUENCES: 113

CORRESPONDENCE ADDRESS:

ADDRESSEE: Arnold, White & Durkee

STREET: P.O. Box 4433

CITY: Houston

STATE: Texas

COUNTRY: USA

ZIP: 77210

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/996.441B

FILING DATE: 18-DEC-1997

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: Kitchell, Barbara S.

REGISTRATION NUMBER: 33,928

REFERENCE/DOCKET NUMBER: MECO:151

TELEPHONE: 512/418-3000

TELEFAX: 512/474-7577

INFORMATION FOR SEQ ID NO: 51:

SEQUENCE CHARACTERISTICS:

LENGTH: 1956 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

FEATURE:

NAME/KEY: CDS

LOCATION: 1..1953

US-08-996-441B-51

Query Match 2.18, Score 20; DB 3; Length 1956;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 ACTTCAACTTGTGCGCATAAA 77  
DB 344 ACTTCAACTTGTGCGCATAAA 325

## RESULT 15

US-08-996-441B-55/C

; Sequence 55, Application US/08996441B

; Patent No. 6023013

; GENERAL INFORMATION:

APPLICANT: English, Leigh H.  
APPLICANT: Brussock, Susan M.  
APPLICANT: Malvar, Thomas M.  
APPLICANT: Bryson, James W.  
APPLICANT: Kulesza, Caroline A.  
APPLICANT: Walters, Frederick S.  
APPLICANT: Slatin, Stephen L.  
APPLICANT: Von Tersch, Michael A.  
APPLICANT: Romano, Charles

TITLE OF INVENTION: INSECT-RESISTANT TRANSGENIC PLANTS

NUMBER OF SEQUENCES: 113

CORRESPONDENCE ADDRESS:

ADDRESSEE: Arnold, White & Durkee

STREET: P.O. Box 4433

CITY: Houston

STATE: Texas

COUNTRY: USA

ZIP: 77210

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/996.441B

FILING DATE: 18-DEC-1997

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: Kitchell, Barbara S.

REGISTRATION NUMBER: 33,928

REFERENCE/DOCKET NUMBER: MECO:151

TELEPHONE: 512/418-3000

TELEFAX: 512/474-7577

INFORMATION FOR SEQ ID NO: 55:

SEQUENCE CHARACTERISTICS:

LENGTH: 1956 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

FEATURE:

NAME/KEY: CDS

LOCATION: 1..1953

US-08-996-441B-55

Query Match 2.18, Score 20; DB 3; Length 1956;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 ACTTCAACTTGTGCGCATAAA 77  
DB 344 ACTTCAACTTGTGCGCATAAA 325

Search completed: June 2, 2003, 05:03:17  
Job time : 64 secs

TELEFAX: (212) 753-6237  
TELEX: 236687  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 918 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Phytolacca insularis  
IMMEDIATE SOURCE:  
CLONE: antiviral protein (PIP)  
US-08-138-636-1

Query Match 2.3%, Score 22; DB 1; Length 918;  
Best Local Similarity 100.0%; Pred. NO. 0.1;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 289 GTGATGGGCTATTCTGATCCCT 310  
Db 277 GTGATGGGCTATTCTGATCCCT 298

## RESULT 12

US-08-319-622A-1  
Sequence 1, Application US/08319622A  
Patent No. 5656466

GENERAL INFORMATION:  
APPLICANT: Moon, Young-Ho  
APPLICANT: Choi, Kyu-Whan  
APPLICANT: Jeon, Hong-Seob  
APPLICANT: Kim, Chul-Hwan  
APPLICANT: Kim, Man-Keun  
TITLE OF INVENTION: Process for Preparing Virus-Resistant  
NUMBER OF SEQUENCES: 4  
TITLE OF INVENTION: Transgenic Plant

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Darby & Darby PC  
STREET: 805 Third Avenue  
CITY: New York  
STATE: New York  
COUNTRY: US  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/319,622A  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Ludwig, S. Peter  
REGISTRATION NUMBER: 25,351  
REFERENCE/DOCKET NUMBER: 0136/0A445  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-527-7700  
TELEFAX: 212-753-6237  
TELEX: 236687

INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 918 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:

ORGANISM: Phytolacca insularis Nakai  
TISSUE TYPE: leaf  
IMMEDIATE SOURCE:  
CLONE: PIP  
US-08-319-622A-1

Query Match 2.3%, Score 22; DB 1; Length 918;  
Best Local Similarity 100.0%; Pred. NO. 0.1;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 289 GTGATGGGCTATTCTGATCCCT 310  
Db 277 GTGATGGGCTATTCTGATCCCT 298

## RESULT 13

US-08-471-564-1  
Sequence 1, Application US/08471564  
Patent No. 5723326

GENERAL INFORMATION:  
APPLICANT: Moon, Young-Ho  
APPLICANT: Jeon, Hong-Seob  
APPLICANT: Choi, Kyu-Whan  
APPLICANT: Lee, Kwan Ho  
APPLICANT: Kim, Man Keun  
TITLE OF INVENTION: NO. 5723326el Genome Coding Phytolacca Antiviral  
NUMBER OF SEQUENCES: 2  
TITLE OF INVENTION: Protein and a Recombinant Expression Vector Therefor  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DARBY & DARBY  
STREET: 805 THIRD AVE.  
CITY: NEW YORK  
STATE: NY  
COUNTRY: USA  
ZIP: 10022

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/471,564  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/210,396  
FILING DATE:  
APPLICATION NUMBER: US 08/138,636  
FILING DATE: 15-OCT-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Ludwig, S. Peter  
REGISTRATION NUMBER: 25,351  
REFERENCE/DOCKET NUMBER: 0136/18818US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 527-7700  
TELEFAX: (212) 753-6237  
TELEX: 236687

INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 918 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Phytolacca insularis  
IMMEDIATE SOURCE:  
CLONE: antiviral protein (PIP)  
US-08-471-564-1

APPLICANT: Tumer, Nilgun  
TITLE OF INVENTION: Biotherapeutic Agents Comprising  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Merchant & Gould  
STREET: 90 South 7th Street, 3100 No. 6146628west Center  
CITY: Minneapolis  
STATE: MN  
COUNTRY: US  
ZIP: 55402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/501.253A  
FILING DATE: 11-JUL-1995  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Kettleberger, Denise M.  
REGISTRATION NUMBER: 33,924  
REFERENCE/DOCKET NUMBER: 600.323US01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 612-332-5300  
TELEFAX: 612-332-9081  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1379 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 225...1163  
US-08-501-253A-1

Query Match 3.1% Score 29: DB 3: Length 1379;  
Best Local Similarity 100.0%; Pred. No. 3e-05;  
Matches 29: Conservative 0: Mismatches 0: Indels 0: Gaps 0:  
Oy 915 TAATCTGGTGATCTATTTCGAAGGATTCT 943  
|||||  
Db 1136 TAATCTGGTGATCTATTTCGAAGGATTCT 1164  
RESULT 10  
PCT-US96-11546-1  
Sequence 1, Application PC/TUS9611546  
GENERAL INFORMATION:  
APPLICANT: Tumer, Nilgun E.  
TITLE OF INVENTION: DNAs Encoding Poxweed Antiviral Protein  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lerner, David, Littenberg, Krumholz & Mentlik  
STREET: 600 South Avenue West  
CITY: Westfield  
STATE: NJ  
COUNTRY: USA  
ZIP: 07090-1497  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/11546  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/500,594  
FILING DATE: 11-JUL-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Foley, Shawn P.  
REFERENCE/DOCKET NUMBER: OCIRS 3.0-039  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 908-654-5000  
TELEFAX: 908-654-7866  
TELEX: 139-175  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1379 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 225...1163  
FEATURE:  
NAME/KEY: mutation  
LOCATION: replace(233, "a")  
FEATURE:  
NAME/KEY: mutation  
LOCATION: replace(349, "g")  
FEATURE:  
NAME/KEY: mutation  
LOCATION: replace(435, "c")  
PCT-US96-11546-1  
Query Match 3.1% Score 29: DB 5: Length 1379;  
Best Local Similarity 100.0%; Pred. No. 3e-05;  
Matches 29: Conservative 0: Mismatches 0: Indels 0: Gaps 0:  
Oy 915 TAATCTGGTGATCTATTTCGAAGGATTCT 943  
|||||  
Db 1136 TAATCTGGTGATCTATTTCGAAGGATTCT 1164  
RESULT 11  
US-08-138-636-1  
Sequence 1, Application US/08138636  
Patent No. 5348865  
GENERAL INFORMATION:  
APPLICANT: Moon, Young-Ho  
APPLICANT: Jeon, Hong-Seon  
APPLICANT: Choi, Kyu-Whan  
APPLICANT: Lee, Kwan Ho  
APPLICANT: Kim, Man Keun  
TITLE OF INVENTION: No. 5348865el Genome Coding Phytolacca Antiviral  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DARBY & DARBY  
STREET: 805 THIRD AVE.  
CITY: NEW YORK  
STATE: NY  
COUNTRY: USA  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/138,636  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Ludwig, S. Peter  
REGISTRATION NUMBER: 25,351  
REFERENCE/DOCKET NUMBER: 0136/08818  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 527-7700

Rest Local Similarity 100.0%; Pred. No. 3e-05; Mismatches 0; Indels 0; Gaps 0;

QY 915 TAATCTTGGTGATCTATTGAAGGATTCT 943  
DB 1136 TAATCTTGGTGATCTATTGAAGGATTCT 1164

RESULT 7  
US-09-005-273-1  
Sequence 1, Application US/09005273  
Patent No. 6137030  
GENERAL INFORMATION:  
APPLICANT: Turner, Nilgun E.  
TITLE OF INVENTION: PAP MUTANTS THAT EXHIBIT ANTI-VIRAL  
TITLE OF INVENTION: AND/OR ANTI-FUNGAL ACTIVITY IN PLANTS  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LERNER, DAVID, LITTENBERG, KRUMHOLZ &  
STREET: 600 South, Avenue West  
CITY: Westfield  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07090  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/005,273  
FILING DATE: 09-JAN-1998  
CLASSIFICATION: 800  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/500,611  
FILING DATE: 11-JUL-1995  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/500,694  
FILING DATE: 11-JUL-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Foley, Shawn P.  
REGISTRATION NUMBER: 33,071  
REFERENCE/DOCKET NUMBER: OCIRS 3.3-034  
TELEPHONE: 908-654-5000  
TELEFAX: 908-654-7866  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1379 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 225..1163  
FEATURE:  
NAME/KEY: sig\_peptide  
LOCATION: 225..290  
US-09-005-273-1

Query Match 3.1%; Score 29; DB 3; Length 1379;  
Best Local Similarity 100.0%; Pred. No. 3e-05;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 915 TAATCTTGGTGATCTATTGAAGGATTCT 943  
DB 1136 TAATCTTGGTGATCTATTGAAGGATTCT 1164

09978274

SRNI

Printed 11/17/2004

RESULT 8  
US-09-005-273-3  
Sequence 3, Application US/09005273  
Patent No. 6137030  
GENERAL INFORMATION:  
APPLICANT: Turner, Nilgun E.  
TITLE OF INVENTION: PAP MUTANTS THAT EXHIBIT ANTI-VIRAL  
TITLE OF INVENTION: AND/OR ANTI-FUNGAL ACTIVITY IN PLANTS  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LERNER, DAVID, LITTENBERG, KRUMHOLZ &  
STREET: 600 South, Avenue West  
CITY: Westfield  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07090  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/005,273  
FILING DATE: 09-JAN-1998  
CLASSIFICATION: 800  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/500,611  
FILING DATE: 11-JUL-1995  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/500,694  
FILING DATE: 11-JUL-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Foley, Shawn P.  
REGISTRATION NUMBER: 33,071  
REFERENCE/DOCKET NUMBER: OCIRS 3.3-034  
TELEPHONE: 908-654-5000  
TELEFAX: 908-654-7866  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1379 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 225..1163  
FEATURE:  
NAME/KEY: sig\_peptide  
LOCATION: 225..290  
US-09-005-273-3

Query Match 3.1%; Score 29; DB 3; Length 1379;  
Best Local Similarity 100.0%; Pred. No. 3e-05;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 915 TAATCTTGGTGATCTATTGAAGGATTCT 943  
DB 1136 TAATCTTGGTGATCTATTGAAGGATTCT 1164

RESULT 9  
US-09-501-253A-1  
Sequence 1, Application US/08501253A  
Patent No. 6146628  
GENERAL INFORMATION:  
APPLICANT: Uckun, Fatih



09978274

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lerner, David, Littenberg, Krumholz & Menlik  
STREET: 600 South Avenue West  
CITY: Westfield  
STATE: NJ  
COUNTRY: USA  
ZIP: 07090-1497

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/500,694  
FILING DATE: 11-JUL-1995  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Foley, Shawn P.  
REFERENCE/DOCKET NUMBER: OCIRS 3.0-039  
TELEPHONE: 908-654-5000  
TELEFAX: 908-654-7866  
TELEX: 139-125

INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1379 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: COS  
LOCATION: 225..1163  
FEATURE:  
NAME/KEY: mutation  
LOCATION: replace(233, "a")  
FEATURE:  
NAME/KEY: mutation  
LOCATION: replace(349, "g")  
FEATURE:  
NAME/KEY: mutation  
LOCATION: replace(435, "c")  
US-08-500-694-1

Query Match 3.1% Score 29; DB 2; Length 1379;  
Best Local Similarity 100.0%; Pred. No. 3e-05;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 915 TAATCTGGTGATCTATTTCAGGATTCT 943  
Db 1136 TAATCTGGTGATCTATTTCAGGATTCT 1164

RESULT 5  
US-07-865-169-1  
Sequence 1, Application US/07865169  
Patent No. 6015940  
GENERAL INFORMATION:  
APPLICANT: Turner, Nilgun E.  
APPLICANT: Lodge, Jennifer K.  
APPLICANT: Kaniewski, Wojciech K.  
TITLE OF INVENTION: Virus Resistant Potato Plants  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co.  
STREET: 700 Chesterfield Parkway No. 6015940th  
CITY: St. Louis  
STATE: Missouri  
COUNTRY: USA  
ZIP: 63198

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

06/02/2003

SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/865,169  
FILING DATE: 19920407  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Hoerner Jr., Dennis R.  
REGISTRATION NUMBER: 30,914  
REFERENCE/DOCKET NUMBER: 38-21(10547)A  
TELEPHONE: (314)537-6099  
TELEFAX: (314)537-6047  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1379 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-07-865-169-1

Query Match 3.1% Score 29; DB 3; Length 1379;  
Best Local Similarity 100.0%; Pred. No. 3e-05;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 915 TAATCTGGTGATCTATTTCAGGATTCT 943  
Db 1136 TAATCTGGTGATCTATTTCAGGATTCT 1164

RESULT 6  
US-07-865-169-2  
Sequence 2, Application US/07865169  
Patent No. 6015940  
GENERAL INFORMATION:  
APPLICANT: Turner, Nilgun E.  
APPLICANT: Lodge, Jennifer K.  
APPLICANT: Kaniewski, Wojciech K.  
TITLE OF INVENTION: Virus Resistant Potato Plants  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co.  
STREET: 700 Chesterfield Parkway No. 6015940th  
CITY: St. Louis  
STATE: Missouri  
COUNTRY: USA  
ZIP: 63198

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/865,169  
FILING DATE: 19920407  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Hoerner Jr., Dennis R.  
REGISTRATION NUMBER: 30,914  
REFERENCE/DOCKET NUMBER: 38-21(10547)A  
TELEPHONE: (314)537-6099  
TELEFAX: (314)537-6047  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1379 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-07-865-169-2

Query Match 3.1% Score 29; DB 3; Length 1379;

ORGANISM: Phytolacca americana  
TISSUE TYPE: Leaf  
US-08-373-858-1

Query Match 3.1% Score 29; DB 1; Length 1195;  
Best Local Similarity 100.0%; Pred. No. 3e-05;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 915 TAATCTTGGTGATCTATTGAAGGATTCT 943  
|||||  
DB 944 TAATCTTGGTGATCTATTGAAGGATTCT 972  
|||||

## RESULT 2

US-08-342-786B-1  
Sequence 1, Application US/08342786B  
Patent No. 5648234

## GENERAL INFORMATION:

APPLICANT: Moon, Young-Ho  
APPLICANT: Jeon, Hong-Seob  
APPLICANT: Choi, Kyu-Whan  
APPLICANT: Lee, Kwan-Ho  
APPLICANT: Kim, Man-Keun  
TITLE OF INVENTION: A No. 5648234el Expression Vector for Phytolacca  
TITLE OF INVENTION: Antiviral Protein  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Darby & Darby  
STREET: 805 Third Ave.  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10022

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/342,786B  
FILING DATE: 21-NOV-1994

## CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/147,024  
FILING DATE: 01-NOV-1993  
CLASSIFICATION: 435

## ATTORNEY/AGENT INFORMATION:

NAME: Ludwig, S. Peter  
REGISTRATION NUMBER: 25,351  
REFERENCE/DOCKET NUMBER: 0136/08862-US1  
TELEPHONE: (212)527-7700  
TELEFAX: (212)753-6237  
TELEX: 236687

## INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:  
LENGTH: 1195 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Phytolacca americana L.  
IMMEDIATE SOURCE:  
CLONE: PAP  
US-08-342-786B-1

Query Match 3.1% Score 29; DB 1; Length 1195;  
Best Local Similarity 100.0%; Pred. No. 3e-05;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 915 TAATCTTGGTGATCTATTGAAGGATTCT 943

06/02/2003

98

DB 944 TAATCTTGGTGATCTATTGAAGGATTCT 972  
|||||

## RESULT 3

US-08-500-611-1  
Sequence 1, Application US/08500611  
Patent No. 5756322

## GENERAL INFORMATION:

APPLICANT: Turner, Nilgun E.  
TITLE OF INVENTION: Pokeweed Antiviral Protein Mutants  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lerner, David, Littenberg, Krumholz & Mentlik  
STREET: 600 South Avenue West  
CITY: Westfield  
STATE: NJ  
COUNTRY: USA  
ZIP: 07090-1497

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/500,611  
FILING DATE: 11-JUL-1995

## CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: Foley, Shawn P.  
REFERENCE/DOCKET NUMBER: OCIRS 3.0-034  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 908-654-5060  
TELEFAX: 908-654-7866  
TELEX: 139-125

## INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:  
LENGTH: 1379 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 225..1163  
FEATURE:  
NAME/KEY: mutation  
LOCATION: replace(233, "a")  
FEATURE:  
NAME/KEY: mutation  
LOCATION: replace(349, "g")  
FEATURE:  
NAME/KEY: mutation  
LOCATION: replace(435, "c")

Query Match 3.1% Score 29; DB 1; Length 1379;  
Best Local Similarity 100.0%; Pred. No. 3e-05;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 915 TAATCTTGGTGATCTATTGAAGGATTCT 943  
|||||  
DB 1136 TAATCTTGGTGATCTATTGAAGGATTCT 1164  
|||||

## RESULT 4

US-08-500-694-1  
Sequence 1, Application US/08500694  
Patent No. 580329

## GENERAL INFORMATION:

APPLICANT: Turner, Nilgun E.  
TITLE OF INVENTION: DNAs Encoding Pokeweed Antiviral Protein  
TITLE OF INVENTION: Mutants  
NUMBER OF SEQUENCES: 2

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 2, 2003, 03:32:48 : Search time 63 seconds  
(without alignments)  
4600.151 Million cell updates/sec

Title: US-09-978-274A-1

Perfect score: 945

Sequence: 1 atgaaggtgatctttagt.....atctattgaagattctaa 945

Scoring table: OLIGO\_NUC

Gapop 60.0, Gapext 60.0

Searched: 441362 seqs, 153338381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued\_Patents\_NA.\*  
1: /cgn2.6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2.6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2.6/ptodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2.6/ptodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2.6/ptodata/1/ina/PCTUS\_COMB.seq.\*  
6: /cgn2.6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	29	3.1	1195	1	US-08-373-858-1
2	29	3.1	1195	1	US-08-342-786B-1
3	29	3.1	1379	2	US-08-500-611-1
4	29	3.1	1379	2	US-08-500-694-1
5	29	3.1	1379	3	US-07-865-169-1
6	29	3.1	1379	3	US-07-865-169-2
7	29	3.1	1379	3	US-09-005-273-1
8	29	3.1	1379	3	US-09-005-273-3
9	29	3.1	1379	3	US-08-501-253A-1
10	29	3.1	1379	5	PCT-US96-11546-1
11	22	2.3	918	1	US-08-138-636-1
12	22	2.3	918	1	US-08-319-622A-1
13	22	2.3	918	1	US-08-471-564-1
14	20	2.1	1956	3	US-08-996-441B-51
15	20	2.1	1956	3	US-08-996-441B-55
16	20	2.1	1956	3	US-08-996-441B-57
17	20	2.1	1956	3	US-08-993-722A-51
18	20	2.1	1956	3	US-08-993-722A-55
19	20	2.1	1956	3	US-08-993-722A-57
20	20	2.1	1956	3	US-08-993-170A-51
21	20	2.1	1956	3	US-08-993-170A-55
22	20	2.1	1956	3	US-08-993-170A-57
23	20	2.1	1956	3	US-08-993-775B-51
24	20	2.1	1956	3	US-08-993-775B-55
25	20	2.1	1956	3	US-08-993-775B-57
26	20	2.1	1959	3	US-08-996-441B-1
27	20	2.1	1959	3	US-08-996-441B-3

c 28	20	2.1	1959	3	US-08-996-441B-5	Sequence 5, Appl
c 29	20	2.1	1959	3	US-08-996-441B-7	Sequence 7, Appl
c 30	20	2.1	1959	3	US-08-996-441B-9	Sequence 9, Appl
c 31	20	2.1	1959	3	US-08-996-441B-11	Sequence 11, Appl
c 32	20	2.1	1959	3	US-08-996-441B-13	Sequence 13, Appl
c 33	20	2.1	1959	3	US-08-996-441B-15	Sequence 15, Appl
c 34	20	2.1	1959	3	US-08-996-441B-17	Sequence 17, Appl
c 35	20	2.1	1959	3	US-08-996-441B-19	Sequence 19, Appl
c 36	20	2.1	1959	3	US-08-996-441B-21	Sequence 21, Appl
c 37	20	2.1	1959	3	US-08-996-441B-23	Sequence 23, Appl
c 38	20	2.1	1959	3	US-08-996-441B-25	Sequence 25, Appl
c 39	20	2.1	1959	3	US-08-996-441B-27	Sequence 27, Appl
c 40	20	2.1	1959	3	US-08-996-441B-29	Sequence 29, Appl
c 41	20	2.1	1959	3	US-08-996-441B-31	Sequence 31, Appl
c 42	20	2.1	1959	3	US-08-996-441B-33	Sequence 33, Appl
c 43	20	2.1	1959	3	US-08-996-441B-35	Sequence 35, Appl
c 44	20	2.1	1959	3	US-08-996-441B-37	Sequence 37, Appl
c 45	20	2.1	1959	3	US-08-996-441B-39	Sequence 39, Appl

ALIGNMENTS

RESULT 1  
US-08-373-858-1  
: Sequence 1, Application US/08373858  
: Patent No. 5633155  
: GENERAL INFORMATION:  
: APPLICANT: Kim, Man-Keun  
: APPLICANT: Lee, Kwan-Ho  
: APPLICANT: Na, Byeong-Kook  
: APPLICANT: Jeong, Han-Seung  
: APPLICANT: Choi, Kyu-Whan  
: APPLICANT: Moon, Young-Ho  
: APPLICANT: Jeon, Hong-Suob  
: TITLE OF INVENTION: Expression vector for Phytolacca  
: TITLE OF INVENTION: Antiviral protein and process for preparing Transgenic  
: TITLE OF INVENTION: Plant Transformed Thereof.  
: NUMBER OF SEQUENCES: 5  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Darby & Darby  
: STREET: 805 Third Avenue  
: CITY: New York  
: STATE: New York  
: COUNTRY: US  
: ZIP: 10022-7513  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patent Release #1.0, Version #1.30  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/373,858  
: FILING DATE: 18-JAN-1995  
: CLASSIFICATION: 435  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Ludwig, Peter  
: REGISTRATION NUMBER: 25,351  
: REFERENCE/DOCKET NUMBER: 0136/17986-US1  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: 212-527-7770  
: TELEFAX: 212-751-6237  
: TELEX: 236687  
: INFORMATION FOR SEQ ID NO: 1:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 1195 base pairs  
: TYPE: nucleic acid  
: STRANDEDNESS: single  
: TOPOLOGY: linear  
: MOLECULE TYPE: cDNA to mRNA  
: HYPOTHETICAL: NO  
: ANTI-SENSE: NO  
: ORIGINAL SOURCE:



NAME: Borun, Michael F.  
REGISTRATION NUMBER: 25,447  
REFERENCE/DOCKET NUMBER: 27129/33199  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6300  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 251:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 332 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 251:  
US-09-765-527-251

Query Match 20.1% Score 325; DB 10; Length 332;  
Best Local Similarity 32.4%; Pred. No. 6e-22;  
Matches 88; Conservative 50; Mismatches 112; Indels 22; Gaps 8;  
QY 15 LIAAPTSTCAINTITFDAGNATINKYATPMESLRNOAKDPKLCYGIPLPDNTSTPK--72  
DB 14 LLAAPAMAGLOTVSFSTKGATYYVNFELNRVCLK-PEGNSHGIGPLLRKKCDOPGKC 72  
QY 73 YLLVLOGANLKITLMLRNNLYVMGYSDPFNGNCKRYHIFNDITSTERTDV-ENTLCS 131  
DB 73 FVLVALSNGOLAEIAIDVTSVYVGYQ-----VNRSYFFKADAPDAAYEGLEFKNT---124  
QY 132 SSSSRVAMSYNLSYPTMEKKAENVSRNOVOLGIGLSSDICKI--SGVDSFPVKTEAF 189  
DB 125 -----IKRLHGGTYPSLE--GEKAYRETTDLGIEPLRIGIKKLDENAIIDYKPTKPTIAS 177  
QY 190 FLVVAIQVSEARFYENQVKTNFNRAFYDPKVINLEEKWKGKISEAH-NAKNGALP 248  
DB 178 SLVVIQVSEARFYENQVKTNFNRAFYDPKVINLEEKWKGKISEAH-NAKNGALP 248  
QY 249 KPLELVDAGKTWIVLRVDEINRDVALLKYN 280  
DB 238 EAVELERANGKYYVTAVDQVAPKIALAKFVD 269

RESULT 14  
US-09-765-527-247  
Sequence 247, Application US/09765527  
Patent No. US20030006638A1  
GENERAL INFORMATION:  
APPLICANT: Better, Marc D.  
TITLE OF INVENTION: Methods for Recombinant Microbial Production of Fusion Proteins and BPI-Derived Peptides  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/765,527  
FILING DATE: 18-Jan-2001  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/621,803  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Borun, Michael F.  
REGISTRATION NUMBER: 25,447  
REFERENCE/DOCKET NUMBER: 27129/33199  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/474-6300  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 247:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 251 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 247:  
US-09-765-527-247

Query Match 19.3% Score 312; DB 10; Length 251;  
Best Local Similarity 32.4%; Pred. No. 6.4e-21;  
Matches 85; Conservative 48; Mismatches 107; Indels 22; Gaps 8;  
QY 25 INTITFDAGNATINKYATPMESLRNOAKDPKLCYGIPLPDNTSTPK--YLLVLOGAN 82  
DB 2 LOTVSFSTKGATYYVNFELNRVCLK-PEGNSHGIGPLLRKKCDOPGKCFCFVLVALSWIN 60  
QY 83 LKTYITLMLRNNLYVMGYSDPFNGNCKRYHIFNDITSTERTDV-ENTLCSSSSSSKVAMSI 141  
DB 61 GOLAEIAIDVTSVYVGYQ-----VNRSYFFKADAPDAAYEGLEFKNT-----IKTKL 107  
QY 142 NYNSLYPTMEKKAENVSRNOVOLGIGLSSDICKI--SGVDSFPVKTEAFLLVVAIQVMS 199  
DB 108 HFGGYSYPSLE--GEKAYRETTDLGIEPLRIGIKKLDENAIIDYKPTKPTIASLLVVAIQVMS 165  
QY 200 EAARKYENQVKTNFNRAFYDPKVINLEEKWKGKISEAH-NAKNGALPDKPILVDAKG 258  
DB 166 EAARTFIENQVKTNFNRAFYDPKVINLEEKWKGKISEAH-NAKNGALPDKPILVDAKG 258  
QY 259 TKWIVLRVDEINRDVALLKYN 280  
DB 226 KKYVTVAVDQVAPKIALAKFVD 247

RESULT 15  
US-09-792-793A-36  
Sequence 36, Application US/09792793A  
Patent No. US20020168370A1  
GENERAL INFORMATION:  
APPLICANT: McDonald, John R.  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING SECONDARY TISSUE DAMAGE  
FILE REFERENCE: 25020-601D  
CURRENT FILING DATE: 2001-02-22  
NUMBER OF SEQ ID NOS: 93  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 36  
LENGTH: 250  
TYPE: PRT  
ORGANISM: Momordica charantia  
US-09-792-793A-36

Query Match 16.1% Score 261; DB 9; Length 250;  
Best Local Similarity 32.6%; Pred. No. 3.3e-16;  
Matches 86; Conservative 41; Mismatches 117; Indels 20; Gaps 10;  
QY 18 APTSTCAINTITFDAGNATINKYATPMESLRNOAKDPKLCYGIPLPDNTSTPKYLLV 76  
DB 1 APT-----LETIASLUNNPT--TYISFITNITKTVADKTEQC-TIOKISKT-FTOKYSVI 52  
QY 77 KLGANLKITLMLRNNLYVMGYSDPFNGNCKRYHIFNDITSTERTDV-ENTLCSSSSK 136  
DB 53 DLIVSSTQKITLADWADLYVLGYSDIAN-NKGRAFFKDYTEA-----VANFFPNTVT 107  
QY 137 VAMSYNLSYPTMEKKAENVSRNOVOLGIGLSSDICKI-SCVDSFPVKTEAFLLVVAIQ 196  
DB 108 NKIKLTFSTGSDLEKNGCLRKDN--PLCJIRLENSIVNIYG-KAGDVKKQAKFFLLAIQ 164

SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: 05/09/765,527  
FILING DATE: 18-Jan-2001  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/621,803  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Borun, Michael F.  
REGISTRATION NUMBER: 25,447  
REFERENCE/DOCKET NUMBER: 27129/33199  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6300  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 259:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 293 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 259:  
US-09-765-527-259

Query Match 20.1% Score 325; DB 10; Length 293;  
Best Local Similarity 32.4% Pred No. 5e-22;  
Matches 88; Conservative 50; Mismatches 112; Indels 22; Gaps 8;  
QY 15 LIAAPTSTCAINTITFDAGNATINKYATPMESLRNOAKPKLKCYGIPMLPDTNSTPK-- 72  
DB 14 LLAQAQAMAGLDTVSFSTKGATITVYVNFNLRLVKLK-PEGNSHGIPLLRKKCDPGKC 72  
QY 73 YLLVKLOGANLKTITLMLRNLNLYGVGSDPFGNKKRYHIFNDITSTERTD-ENTLCS 131  
DB 73 FVLVLSNDNGQLAEIAIDVTSVYVVGQ-----VNRSYFFKADDAAYEGLFKNT--- 124  
QY 132 SSSSRVAMSYNLSYPTMEKKAENSRNOVQIGIQLSSDICKI--SGVDSFPVKTEAF 189  
DB 125 -----IKRLHFGGTPSLE--GEKAYRETTDLGIRPLRIGIKKLDENATDNYKPTETAS 177  
QY 190 FLVVAIQWSEARFKYIENQVKTENRATYDPKPVINLEKKGKISEATH-NAKNGCALP 248  
DB 178 SLVVQIQWSEARFTFIENQIRNFQIRPANNITISLENKWKLSFQIRTSGANGMFS 237  
QY 249 KPLELDVAKGKTWILVRVDEINRQVALLKYVN 280  
DB 238 EAVELERANGKKYVYVAVDQVQPKIALLEKVD 269

RESULT 12  
US-09-765-527-253  
Sequence 253, Application US/09765527  
Patent No. US20020006638A1  
GENERAL INFORMATION:  
APPLICANT: Better, Marc D.  
TITLE OF INVENTION: Methods for Recombinant Microbial Production of Fusion Proteins and BPI-Derived Peptides  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/765,527  
FILING DATE: 18-Jan-2001

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/621,803  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Borun, Michael F.  
REGISTRATION NUMBER: 25,447  
REFERENCE/DOCKET NUMBER: 27129/33199  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6300  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 253:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 309 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 253:  
US-09-765-527-253

Query Match 20.1% Score 325; DB 10; Length 309;  
Best Local Similarity 32.4% Pred No. 5.4e-22;  
Matches 88; Conservative 50; Mismatches 112; Indels 22; Gaps 8;  
QY 15 LIAAPTSTCAINTITFDAGNATINKYATPMESLRNOAKPKLKCYGIPMLPDTNSTPK-- 72  
DB 14 LLAQAQAMAGLDTVSFSTKGATITVYVNFNLRLVKLK-PEGNSHGIPLLRKKCDPGKC 72  
QY 73 YLLVKLOGANLKTITLMLRNLNLYGVGSDPFGNKKRYHIFNDITSTERTD-ENTLCS 131  
DB 73 FVLVLSNDNGQLAEIAIDVTSVYVVGQ-----VNRSYFFKADDAAYEGLFKNT--- 124  
QY 132 SSSSRVAMSYNLSYPTMEKKAENSRNOVQIGIQLSSDICKI--SGVDSFPVKTEAF 189  
DB 125 -----IKRLHFGGTPSLE--GEKAYRETTDLGIRPLRIGIKKLDENATDNYKPTETAS 177  
QY 190 FLVVAIQWSEARFKYIENQVKTENRATYDPKPVINLEKKGKISEATH-NAKNGCALP 248  
DB 178 SLVVQIQWSEARFTFIENQIRNFQIRPANNITISLENKWKLSFQIRTSGANGMFS 237  
QY 249 KPLELDVAKGKTWILVRVDEINRQVALLKYVN 280  
DB 238 EAVELERANGKKYVYVAVDQVQPKIALLEKVD 269

RESULT 13  
US-09-765-527-251  
Sequence 251, Application US/09765527  
Patent No. US20020006638A1  
GENERAL INFORMATION:  
APPLICANT: Better, Marc D.  
TITLE OF INVENTION: Methods for Recombinant Microbial Production of Fusion Proteins and BPI-Derived Peptides  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/765,527  
FILING DATE: 18-Jan-2001  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/621,803  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:

LENGTH: 275  
TYPE: PRT  
ORGANISM: Saponaria officinalis  
US-09-792-793A-35

Query Match 21.2% Score 342; DB 9; Length 275;  
Best Local Similarity 35.9% Pred. No. 1.6e-23;  
Matches 94; Conservative 51; Mismatches 97; Indels 20; Gaps 9;

Qy 25 INTITDAGNATINKYATFMSLRNQAOKPKLKYGIPM-LPDTNSTPKYLLVKGANL 83  
Db 1 VTSITDLVNPTAGQYSSFDKIRNNVKDPLNKYGGTDIAVIGPPSKKEFLRINFQSSR- 59  
Qy 84 KTTTLMLRNNLYVMGYSDPFNGNKR-YHIFNDITSTERTDVENTLCSSSSSRVAMSN 142  
Db 60 GTVSLGLKRDNLVYVAYLAMDNTNVNRYAYRSEITSAST-ALFPEATTAQNAKALE 115  
Qy 143 YNSLYPTMEKKAEV-NSRNOVOLGIGIOLSSDICKISGVDSFPVKTEAFFLLVATOMV 198  
Db 116 YTEDYOSIEKNAQTQGDOSRKELGLDLSSTMEAVN-KKARVVKDEARFLIATQMT 174  
Qy 199 SEARFKYIENOVKTNFRNRYFDPKVINLEEKWGISIAIH-NAKNGALPKPLELVDK 257  
Db 175 AEAAFRYIQLVKNFKNSENKVIQFEVNNKKISTAIYGDKNGVFNKDYDFGFGK 234  
Qy 258 GTKWIVLRVDEINRDVALLKYV 279  
Db 235 -----VRQVKDL--QMGLLMYL 249

RESULT 9  
US-09-792-793A-79

Sequence 79, Application US/09792793A  
Patent No. US20020168370A1  
GENERAL INFORMATION:  
APPLICANT: McDonald, John R.  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING SECONDARY TISSUE DAMAGE AND  
TITLE OF INVENTION: OTHER INFLAMMATORY CONDITIONS AND DISORDERS  
FILE REFERENCE: 25020-601D  
CURRENT APPLICATION NUMBER: US/09/792,793A  
CURRENT FILING DATE: 2001-02-22  
NUMBER OF SEQ ID NOS: 93  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO 79  
LENGTH: 327  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Chemokine-toxin fusion protein  
OTHER INFORMATION: 1-Beta-AM-SAPORIN  
US-09-792-793A-79

Query Match 21.2% Score 342; DB 9; Length 327;  
Best Local Similarity 35.9% Pred. No. 1.6e-23;  
Matches 94; Conservative 51; Mismatches 97; Indels 20; Gaps 9;

Qy 25 INTITDAGNATINKYATFMSLRNQAOKPKLKYGIPM-LPDTNSTPKYLLVKGANL 83  
Db 75 VTSITDLVNPTAGQYSSFDKIRNNVKDPLNKYGGTDIAVIGPPSKKEFLRINFQSSR- 133  
Qy 84 KTTTLMLRNNLYVMGYSDPFNGNKR-YHIFNDITSTERTDVENTLCSSSSSRVAMSN 142  
Db 134 GTVSLGLKRDNLVYVAYLAMDNTNVNRYAYRSEITSAST-ALFPEATTAQNAKALE 189  
Qy 143 YNSLYPTMEKKAEV-NSRNOVOLGIGIOLSSDICKISGVDSFPVKTEAFFLLVATOMV 198  
Db 190 YTEDYOSIEKNAQTQGDOSRKELGLDLSSTMEAVN-KKARVVKDEARFLIATQMT 248  
Qy 199 SEARFKYIENOVKTNFRNRYFDPKVINLEEKWGISIAIH-NAKNGALPKPLELVDK 257  
Db 249 AEAAFRYIQLVKNFKNSENKVIQFEVNNKKISTAIYGDKNGVFNKDYDFGFGK 308

Qy 258 GTKWIVLRVDEINRDVALLKYV 279  
Db 309 -----VRQVKDL--QMGLLMYL 323

RESULT 10

US-09-792-793A-76  
Sequence 76, Application US/09792793A  
Patent No. US20020168370A1  
GENERAL INFORMATION:  
APPLICANT: McDonald, John R.  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING SECONDARY TISSUE DAMAGE  
TITLE OF INVENTION: OTHER INFLAMMATORY CONDITIONS AND DISORDERS  
FILE REFERENCE: 25020-601D  
CURRENT APPLICATION NUMBER: US/09/792,793A  
CURRENT FILING DATE: 2001-02-22  
NUMBER OF SEQ ID NOS: 93  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO 76  
LENGTH: 332  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Chemokine-toxin fusion  
OTHER INFORMATION: MCH3-AM-SAPORIN  
US-09-792-793A-76

Query Match 21.2% Score 342; DB 9; Length 332;  
Best Local Similarity 35.9% Pred. No. 1.6e-23;  
Matches 94; Conservative 51; Mismatches 97; Indels 20; Gaps 9;

Qy 25 INTITDAGNATINKYATFMSLRNQAOKPKLKYGIPM-LPDTNSTPKYLLVKGANL 83  
Db 80 VTSITDLVNPTAGQYSSFDKIRNNVKDPLNKYGGTDIAVIGPPSKKEFLRINFQSSR- 138  
Qy 84 KTTTLMLRNNLYVMGYSDPFNGNKR-YHIFNDITSTERTDVENTLCSSSSSRVAMSN 142  
Db 139 GTVSLGLKRDNLVYVAYLAMDNTNVNRYAYRSEITSAST-ALFPEATTAQNAKALE 194  
Qy 143 YNSLYPTMEKKAEV-NSRNOVOLGIGIOLSSDICKISGVDSFPVKTEAFFLLVATOMV 198  
Db 195 YTEDYOSIEKNAQTQGDOSRKELGLDLSSTMEAVN-KKARVVKDEARFLIATQMT 254  
Qy 199 SEARFKYIENOVKTNFRNRYFDPKVINLEEKWGISIAIH-NAKNGALPKPLELVDK 257  
Db 254 AEAAFRYIQLVKNFKNSENKVIQFEVNNKKISTAIYGDKNGVFNKDYDFGFGK 313  
Qy 258 GTKWIVLRVDEINRDVALLKYV 279  
Db 314 -----VRQVKDL--QMGLLMYL 328

RESULT 11  
US-09-765-527-259

Sequence 259, Application US/09765527  
Patent No. US2002008638A1  
GENERAL INFORMATION:  
APPLICANT: Better, Marc D.  
TITLE OF INVENTION: Methods for Recombinant Microbial Production of  
Fusion Proteins and BPI-Derived Peptides  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borum  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

06/02/2003

Db 185 TANQKALEYTTEDYOSIEKNAQITOGDOSRKELGIGLIDLLSTSMYAVN-KKARVVKDEARF 243  
Qy 191 LLVATQYVSEAREKYTENQVTKNFNRAFYDPKVINIEKWKGISEAIBH-NAKNGALPK 249  
Db 244 LLJAJQMTAEARFYQNLVTKFNPKNFSENKVIQFELVNNKKISTATYGDANKGVNFK 303  
Qy 250 PUELVDAKGTWVLRVDEINRQVALLKYV 279  
Db 304 DYDFGFGK-----VRQVKDL--QMGLLMYL 326

RESULT 7

US-09-792-793A-85  
; Sequence 85, Application US/09792793A  
; Patent No. US20020168370A1  
; GENERAL INFORMATION:  
; APPLICANT: McDonald, John R.  
; APPLICANT: Coggin, Philip  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING SECONDARY TISSUE DAMAGE  
; FILE REFERENCE: 25020-601D  
; CURRENT APPLICATION NUMBER: US/09/792,793A  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 93  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 85  
; LENGTH: 254  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Methionine-Saporin fusion  
US-09-792-793A-85

Query Match

Best Local Similarity 21.2%; Score 342; DB 9; Length 254;

Matches 94; Conservative 51; Mismatches 97; Indels 20; Gaps 9;

RESULT 8

US-09-792-793A-85  
; Sequence 85, Application US/09792793A  
; Patent No. US20020168370A1  
; GENERAL INFORMATION:  
; APPLICANT: McDonald, John R.  
; APPLICANT: Coggin, Philip  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING SECONDARY TISSUE DAMAGE  
; FILE REFERENCE: 25020-601D  
; CURRENT APPLICATION NUMBER: US/09/792,793A  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 93  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 85  
; LENGTH: 254  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Methionine-Saporin fusion  
US-09-792-793A-85

Query Match

Best Local Similarity 21.2%; Score 342; DB 9; Length 254;

Matches 94; Conservative 51; Mismatches 97; Indels 20; Gaps 9;

RESULT 9

US-09-792-793A-85  
; Sequence 85, Application US/09792793A  
; Patent No. US20020168370A1  
; GENERAL INFORMATION:  
; APPLICANT: McDonald, John R.  
; APPLICANT: Coggin, Philip  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING SECONDARY TISSUE DAMAGE  
; FILE REFERENCE: 25020-601D  
; CURRENT APPLICATION NUMBER: US/09/792,793A  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 93  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 85  
; LENGTH: 254  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Methionine-Saporin fusion  
US-09-792-793A-85

09978274

SRNT

Printed 11/17/2004



09978274

OY 241 NAKNGALPKPLRLVDAGTKWIVLRVDEINRDVALLKYVNGTCOTTQYQNAFMSQVILSTY 300  
DB 241 NAKNGALPKPLELVDAKGTWIVLRVDEINRDVALLKYVNGTCOTTQYQNAFMSQVILSTY 300  
OY 301 YNYMSNLGDLFEFG 314  
DB 301 YNYMSNLGDLFEFG 314

RESULT 2  
US-09-978-274A-4  
: Sequence 4, Application US/09978274A  
: Patent No. US20020116737A1  
: GENERAL INFORMATION:  
: APPLICANT: Thomas, Christopher  
: APPLICANT: McPherson, Michael  
: APPLICANT: Atkinson, Howard  
: APPLICANT: Neelam, Anil  
: TITLE OF INVENTION: PLANT CELL DEATH SYSTEM  
: FILE REFERENCE: 9341-028  
: CURRENT APPLICATION NUMBER: US/09/978, 274A  
: CURRENT FILING DATE: 2001-10-15  
: PRIOR APPLICATION NUMBER: 0025225.4  
: PRIOR FILING DATE: 2000-10-14  
: NUMBER OF SEQ ID NOS: 32  
: SOFTWARE: PatentIn version 3.1  
: SEQ ID NO 4  
: LENGTH: 263  
: TYPE: PRT  
: ORGANISM: Phytolacca americana  
US-09-978-274A-4

Query Match 83.6%; Score 1352; DB 10; Length 263;  
Best Local Similarity 100.0%; Pred. No. 6e-117;  
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 25 INTITFDAGNATINKYATFMSLRNOAKDPKLCYGPIMPLDPTNSTPKYLLVKLOGANLK 84  
DB 2 INTITFDAGNATINKYATFMSLRNOAKDPKLCYGPIMPLDPTNSTPKYLLVKLOGANLK 61  
OY 85 TITLMLRRNNLYVMGYSDPFNGKNCRYHIFNDITSTERTDVENTLCSSSSSRVAMSINYN 144  
DB 62 TITLMLRRNNLYVMGYSDPFNGKNCRYHIFNDITSTERTDVENTLCSSSSSRVAMSINYN 121  
OY 145 SLYPTMEKKAENVSRNOVOLGQILSSDGIKISGVDSFPVKTEAFFLLVAIQMVSEAAAF 204  
DB 122 SLYPTMEKKAENVSRNOVOLGQILSSDGIKISGVDSFPVKTEAFFLLVAIQMVSEAAAF 181  
OY 205 KYIENQVKTNFNRAFYDPKPKYINLEEKWKXISEATHNAKNGALPKPLELVDAKTKWIVL 264  
DB 182 KYIENQVKTNFNRAFYDPKPKYINLEEKWKXISEATHNAKNGALPKPLELVDAKTKWIVL 241  
OY 265 RVDEINRDVALLKYVNGTCOTT 286  
DB 242 RVDEINRDVALLKYVNGTCOTT 263

RESULT 3  
US-09-978-274A-6  
: Sequence 6, Application US/09978274A  
: Patent No. US20020116737A1  
: GENERAL INFORMATION:  
: APPLICANT: Thomas, Christopher  
: APPLICANT: McPherson, Michael  
: APPLICANT: Atkinson, Howard  
: APPLICANT: Neelam, Anil  
: TITLE OF INVENTION: PLANT CELL DEATH SYSTEM  
: FILE REFERENCE: 9341-028  
: CURRENT APPLICATION NUMBER: US/09/978, 274A  
: CURRENT FILING DATE: 2001-10-15  
: PRIOR APPLICATION NUMBER: 0025225.4  
: PRIOR FILING DATE: 2000-10-14

06/02/2003

: NUMBER OF SEQ ID NOS: 32  
: SOFTWARE: PatentIn version 3.1  
: SEQ ID NO 6  
: LENGTH: 154  
: TYPE: PRT  
: ORGANISM: Phytolacca americana  
US-09-978-274A-6

Query Match 48.7%; Score 788; DB 10; Length 154;  
Best Local Similarity 100.0%; Pred. No. 3.5e-65;  
Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 25 INTITFDAGNATINKYATFMSLRNOAKDPKLCYGPIMPLDPTNSTPKYLLVKLOGANLK 84  
DB 2 INTITFDAGNATINKYATFMSLRNOAKDPKLCYGPIMPLDPTNSTPKYLLVKLOGANLK 61  
OY 85 TITLMLRRNNLYVMGYSDPFNGKNCRYHIFNDITSTERTDVENTLCSSSSSRVAMSINYN 144  
DB 62 TITLMLRRNNLYVMGYSDPFNGKNCRYHIFNDITSTERTDVENTLCSSSSSRVAMSINYN 121  
OY 145 SLYPTMEKKAENVSRNOVOLGQILSSDGIKIS 177  
DB 122 SLYPTMEKKAENVSRNOVOLGQILSSDGIKIS 154

RESULT 4  
US-09-978-274A-8  
: Sequence 8, Application US/09978274A  
: Patent No. US20020116737A1  
: GENERAL INFORMATION:  
: APPLICANT: Thomas, Christopher  
: APPLICANT: McPherson, Michael  
: APPLICANT: Atkinson, Howard  
: APPLICANT: Neelam, Anil  
: TITLE OF INVENTION: PLANT CELL DEATH SYSTEM  
: FILE REFERENCE: 9341-028  
: CURRENT APPLICATION NUMBER: US/09/978, 274A  
: CURRENT FILING DATE: 2001-10-15  
: PRIOR APPLICATION NUMBER: 0025225.4  
: PRIOR FILING DATE: 2000-10-14  
: NUMBER OF SEQ ID NOS: 32  
: SOFTWARE: PatentIn version 3.1  
: SEQ ID NO 8  
: LENGTH: 110  
: TYPE: PRT  
: ORGANISM: Phytolacca americana  
US-09-978-274A-8

Query Match 34.9%; Score 564; DB 10; Length 110;  
Best Local Similarity 100.0%; Pred. No. 1.1e-44;  
Matches 109; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 178 GVDSPVKTEAFFLLVAIQMVSEAAAFKYIENQVKTNFNRAFYDPKPKYINLEEKWKXISE 217  
DB 2 GVDSPVKTEAFFLLVAIQMVSEAAAFKYIENQVKTNFNRAFYDPKPKYINLEEKWKXISE 61  
OY 238 AHNKNGALPKPLELVDAKTKWIVLRVDEINRDVALLKYVNGTCOTT 286  
DB 62 AHNKNGALPKPLELVDAKTKWIVLRVDEINRDVALLKYVNGTCOTT 110

RESULT 5  
US-09-792-793A-73  
: Sequence 73, Application US/09792793A  
: Patent No. US2002016370A1  
: GENERAL INFORMATION:  
: APPLICANT: McDonald, John R.  
: APPLICANT: Cozzins, Philip  
: TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING SECONDARY TISSUE DAMAGE  
: FILE REFERENCE: 2502-601D  
: CURRENT APPLICATION NUMBER: US/09/792,793A  
: CURRENT FILING DATE: 2001-02-22

SRNT

Printed 11/17/2004

GenCore version 5.1.4.p2\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

CM protein - protein search, using sw model

Run on: May 28, 2003, 09:58:20 ; Search time 58 seconds  
(without alignments)  
536,900 Million cell updates/sec

Title: US-09-978-274A-2

Perfect score: 1617

Sequence: 1 MKVMLVVVVTLIAWLIAPT.....VILSTYNYMNLGOLFEGF 314

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 375593 seqs, 99172665 residues

Total number of hits satisfying chosen parameters: 375593

Minimum DH seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

Published\_Applications\_AA:

- 1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US05\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1617	100.0	314	10	US-09-978-274A-2
2	1352	83.6	263	10	US-09-978-274A-4
3	788	48.7	154	10	US-09-978-274A-6
4	564	34.9	110	10	US-09-978-274A-8
5	346	21.4	332	9	US-09-792-793A-73
6	343.5	21.2	330	9	US-09-792-793A-82
7	342	21.2	254	9	US-09-792-793A-85
8	342	21.2	275	9	US-09-792-793A-35
9	342	21.2	327	9	US-09-792-793A-79
10	342	21.2	332	9	US-09-792-793A-76
11	325	20.1	293	10	US-09-765-527-259
12	325	20.1	309	10	US-09-765-527-253
13	325	20.1	332	10	US-09-765-527-251
14	312	19.3	251	10	US-09-765-527-247
15	261	16.1	250	9	US-09-792-793A-36
16	240	14.8	247	9	US-09-792-793A-39
17	225	13.9	247	9	US-09-792-793A-34
18	204	12.6	252	10	US-09-347-064-2
19	204	12.6	252	10	US-09-347-064-8

20	130.5	8.1	315	10	US-09-334-477-2	Sequence 2, Appl
21	130.5	8.1	323	10	US-09-334-477-21	Sequence 21, Appl
22	130	8.0	323	9	US-09-792-793A-80	Sequence 80, Appl
23	130	8.0	325	9	US-09-792-793A-81	Sequence 81, Appl
24	125.5	7.8	325	9	US-09-792-793A-74	Sequence 74, Appl
25	125.5	7.8	327	9	US-09-792-793A-75	Sequence 75, Appl
26	124	7.7	325	9	US-09-792-793A-71	Sequence 71, Appl
27	124	7.7	327	9	US-09-792-793A-72	Sequence 72, Appl
28	123	7.6	326	10	US-09-334-477-37	Sequence 37, Appl
29	123	7.6	650	10	US-09-334-477-47	Sequence 47, Appl
30	123	7.6	708	10	US-09-334-477-33	Sequence 33, Appl
31	122	7.5	320	9	US-09-792-793A-77	Sequence 77, Appl
32	122	7.5	322	9	US-09-792-793A-78	Sequence 78, Appl
33	121	7.5	319	9	US-09-792-793A-38	Sequence 38, Appl
34	121	7.5	319	9	US-09-870-759-28	Sequence 28, Appl
35	120	7.4	247	9	US-09-792-793A-83	Sequence 83, Appl
36	120	7.4	249	9	US-09-792-793A-84	Sequence 84, Appl
37	120	7.4	293	9	US-09-792-793A-37	Sequence 37, Appl
38	119.5	7.4	694	10	US-09-334-477-49	Sequence 49, Appl
39	112.5	7.0	318	10	US-09-334-477-6	Sequence 6, Appl
40	112.5	7.0	326	10	US-09-334-477-25	Sequence 25, Appl
41	111	6.9	329	10	US-09-334-477-39	Sequence 39, Appl
42	111	6.9	711	10	US-09-334-477-35	Sequence 35, Appl
43	95.5	5.9	1447	10	US-09-797-097-2	Sequence 2, Appl
44	93	5.8	1974	9	US-09-895-913A-12	Sequence 12, Appl
45	89.5	5.5	930	10	US-09-815-242-10779	Sequence 10779, A

ALIGNMENTS

RESULT 1

US-09-978-274A-2  
: Sequence 2, Application US/09978274A  
: Patent No. US20020116737A1  
: GENERAL INFORMATION:  
: APPLICANT: Thomas, Christopher  
: APPLICANT: McPherson, Michael  
: APPLICANT: Atkinson, Howard  
: APPLICANT: Neelam, Anil  
: TITLE OF INVENTION: PLANT CELL DEATH SYSTEM  
: FILE REFERENCE: 9341-028  
: CURRENT APPLICATION NUMBER: US/09/978, 274A  
: PRIOR FILING DATE: 2001-10-15  
: PRIOR APPLICATION NUMBER: 0025225.4  
: PRIOR FILING DATE: 2000-10-14  
: NUMBER OF SEQ ID NOS: 32  
: SOFTWARE: PatentIn version 3.1  
: SEQ ID NO 2  
: LENGTH: 314  
: TYPE: PKT  
: ORGANISM: Phytolacca americana  
US-09-978-274A-2

Query Match	100.0%	Score 1617:	DB 10:	Length 314:
Best Local Similarity	100.0%:	Pred. No. 2.6e-141:		
Matches: 314:	Conservative 0:	Mismatches 0:	Indels 0:	Gaps 0:
Oy	1	MKVMLVVVVTLIAWLIAPTSTCAINTITFDAGNATINKYATFMESLRNOAKDPKLCYGG	60	
Db	1	MKVMLVVVVTLIAWLIAPTSTCAINTITFDAGNATINKYATFMESLRNOAKDPKLCYGG	60	
Oy	61	TPMLPDTNSTPKYLLVKLOGANIKTITLIMLRNNLYVMCYSDPPFNGKCRVITFMFIST	120	
Db	61	TPMLPDTNSTPKYLLVKLOGANIKTITLIMLRNNLYVMCYSDPPFNGKCRVITFMFIST	120	
Oy	121	ERTDVNTLCSSSSSKVAMSYNLSLYPTMEKKAQVNSRQVGLGQIQLSSDICKISQVQ	180	
Db	121	ERTDVNTLCSSSSSKVAMSYNLSLYPTMEKKAQVNSRQVGLGQIQLSSDICKISQVQ	180	
Oy	181	SPFVKTEAFFLLVATQWSEAKPKYIENOVKTNPKNFAPYPOPKVINLEEKWKISIAH	240	
Db	181	SPFVKTEAFFLLVATQWSEAKPKYIENOVKTNPKNFAPYPOPKVINLEEKWKISIAH	240	

240 RVDEIKPDVGLLNYVNGTCQAT 261

RESULT 15

0009-610-838-9

Sequence 9, Application US/09610838

Patent No. 6376217

GENERAL INFORMATION:

APPLICANT: Better, Marc D.

APPLICANT: Carroll, Stephen F.

APPLICANT: Sudhika, Gary M.

TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating

TITLE OF INVENTION: Proteins

NUMBER OF SEQUENCES: 173

CORRESPONDENCE ADDRESS:

ADDRESSEE: McAndrews, Heid &amp; Malloy, Ltd.

STREET: 500 West Madison Street, 34th floor

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60661

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/610,838

FILING DATE: 06-JUL-2000

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/136,389

FILING DATE: 18-AUG-1998

APPLICATION NUMBER: 08/646,360

FILING DATE: 13-MAY-1996

APPLICATION NUMBER: PCT/US94/05348

FILING DATE: 12-MAY-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/064,691

FILING DATE: 12-MAY-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/988,430

FILING DATE: 09-DEC-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/901,707

FILING DATE: 19-JUN-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/787,567

FILING DATE: 04-NOV-1991

ATTORNEY/AGENT INFORMATION:

NAME: McNicholas, Janet M.

REGISTRATION NUMBER: 32,918

REFERENCE/DOCKET NUMBER: 200-70-P4

TELEPHONE: 312/707-8889

TELEFAX: 312/707-9155

TELEX: 650 388-1248

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 261 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

0009-610-838-9

Query Match

67.88; Score 1096.5; DB 4: Length 261;

Best Local Similarity

82.88; Pred. No. 1.5e-101;

Matches 217; Conservative

14; Mismatches 30; Indels 1; Caps 1;

25 INTTFDAGNATINKYATPMESLRNQAQPKLCYGIPLPDNTNSTPKYLVLKLGANLK 84

1 INTTFDAGNATINKYATPMESLRNQAQPKLCYGIPLPDNTNSTPKYLVLKLGASLK 60

Printed

QY 85 TITLMLRRNNLYVMGYSDPENGKCRHYHFNIDITSTERTDVENTLSSSSSRVAMSINTN 144  
|||||  
DB 61 TITLMLRRNNLYVMGYSDPYD-NKCRHYHFNIDIKGTESDVENTLCPSSNPVAKPYN 119  
|||||  
QY 145 SLYPTHEKKAENVNRKQVGLGIGLSSDICKISGVDSFPVKTEAFFLVAIONVSEAKP 204  
|||||  
DB 120 GLYPTLEKKAGVTSRNEVGLGIGLSSDICKISGVDSFPVKTEAFFLVAIONVSEAKP 179  
|||||  
QY 205 KYTENOVKTNFNEAFYDPKVINLEKKWCKISPAHNKNGALPKPLELVDAKGTWKIVL 264  
|||||  
DB 180 KYTENOVKTNFNKFDFNDKVLPLERNWCKISTAHNSKNGALPKPLELNKNAICTWKIVL 239  
|||||  
QY 265 RVDEINRQVALIKYVNGTCQAT 286  
|||||  
DB 240 RVDEIKPDVGLLNYVNGTCQAT 261  
|||||

Search completed: May 28, 2003, 09:59:41  
Job time : 31 secs

SRNT

0009-610-838-9

Printed

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/839,765  
FILING DATE: 15-APR-1997  
CLASSIFICATION: 530.  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/425,336  
FILING DATE: 18-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/064,691  
FILING DATE: 12-MAY-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/988,430  
FILING DATE: 09-DEC-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: McNicholas, Janet M.  
REGISTRATION NUMBER: 32,918  
REFERENCE/DOCKET NUMBER: 110220US09/200-70.P3.C3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/707-8889  
TELEFAX: 312/707-9155  
TELEX: 650 388-1248  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 261 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-839-765-9

Query Match 67.8% Score 1096.5; DB 4; Length 261;  
Best Local Similarity 82.8% Pred. No. 1.5e-101;  
Matches 217; Conservative 14; Mismatches 30; Indels 1; Gaps 1;  
2y 25 INITFDAGNATINKYATFMSLRNOAKDPKLCYGPMLPDTNSTPKYLLVKQGANLK 84  
3b 1 INITFDAGNATINKYATFMSLRNEAKDPKLCYGPMLPDTNSTPKYLLVKQGANLK 60  
3y 85 TITLMLRRNLNLYVMGYSDPFGNCKRYHIFNDITSTERTDVENTLCSSSSSRVAMSINYN 144  
3b 61 TITLMLRRNLNLYVMGYSDPYD-NKCRYHIFNDIKGTEYSDVENTLCPSSNPRVAKPINYN 119  
3y 145 SLYPTMEKKAENSVNRNOVLGQILSSDICKISGVDSFPVKTEAFELLVAIQVSAARF 204  
3b 120 GLYPTLEKKAGVTSRNEVQLGQILSSDICKISGVDSFPVKTEAFELLVAIQVSAARF 179  
3y 205 KYIENOVKTNFNRAFPDPKVINLEEKWKGLKISFAINAKNGALPKPLVLDAGTKWIVL 264  
3b 180 KYIENOVKTNFNRFSPDNKVLQLEENWGLKISTAIHNSKNGALPKPLVLDAGTKWIVL 239  
3y 265 RVDEINRDVALKYVNGTCQTT 286  
3b 240 RVDEIKPDVGLNLYVNGTCOAT 261

Sequence 9, Application US/09136389  
Sequence No. 6146850  
GENERAL INFORMATION:  
APPLICANT: Better, Marc D.  
APPLICANT: Carroll, Stephen F.  
APPLICANT: Studnicka, Gary M.  
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating

TITLE OF INVENTION: Proteins  
NUMBER OF SEQUENCES: 171  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
STREET: 500 West Madison Street, 34th floor  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60661  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/136,389  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/646,360  
FILING DATE: 13-MAY-1996  
APPLICATION NUMBER: PCT/US94/05348  
FILING DATE: 12-MAY-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/064,691  
FILING DATE: 12-MAY-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/988,430  
FILING DATE: 09-DEC-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: McNicholas, Janet M.  
REGISTRATION NUMBER: 32,918  
REFERENCE/DOCKET NUMBER: 200-70.P4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/707-8889  
TELEFAX: 312/707-9155  
TELEX: 650 388-1248  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 261 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-136-389-9

Query Match 67.8% Score 1096.5; DB 4; Length 261;  
Best Local Similarity 82.8% Pred. No. 1.5e-101;  
Matches 217; Conservative 14; Mismatches 30; Indels 1; Gaps 1;  
3y 25 INITFDAGNATINKYATFMSLRNOAKDPKLCYGPMLPDTNSTPKYLLVKQGANLK 84  
3b 1 INITFDAGNATINKYATFMSLRNEAKDPKLCYGPMLPDTNSTPKYLLVKQGANLK 60  
3y 85 TITLMLRRNLNLYVMGYSDPFGNCKRYHIFNDITSTERTDVENTLCSSSSSRVAMSINYN 144  
3b 61 TITLMLRRNLNLYVMGYSDPYD-NKCRYHIFNDIKGTEYSDVENTLCPSSNPRVAKPINYN 119  
3y 145 SLYPTMEKKAENSVNRNOVLGQILSSDICKISGVDSFPVKTEAFELLVAIQVSAARF 204  
3b 120 GLYPTLEKKAGVTSRNEVQLGQILSSDICKISGVDSFPVKTEAFELLVAIQVSAARF 179  
3y 205 KYIENOVKTNFNRAFPDPKVINLEEKWKGLKISFAINAKNGALPKPLVLDAGTKWIVL 264  
3b 180 KYIENOVKTNFNRFSPDNKVLQLEENWGLKISTAIHNSKNGALPKPLVLDAGTKWIVL 239  
3y 265 RVDEINRDVALKYVNGTCQTT 286  
3b 240 RVDEIKPDVGLNLYVNGTCOAT 261

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: McNicholas, Janet M.  
REGISTRATION NUMBER: 32,918  
REFERENCE/DOCKET NUMBER: 11022US07/200-70.P3.C2A  
TELEPHONE: 312/707-8889  
TELEFAX: 312/707-9155  
TELEX: 650 388-1248  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 261 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-477-484B-9

Query Match 67.8%; Score 1096.5; DB 1; Length 261;  
Best Local Similarity 82.8%; Pred. No. 1.5e-101;  
Matches 217; Conservative 14; Mismatches 30; Indels 1; Gaps 1;  
Y 25 INTITFDAGNATINKYATPMESLRNOAKDPKLCYGPMLPDTNSTPKYLLVKLOGANLK 84  
b 1 INTITFDAGNATINKYATPMESLRNEAKDPSLKCYGIPMLPNTNSTIKYLLVKLOGASLK 60  
Y 85 TITLMLRRNLNLYVMGYSDFPENGKCRHYFNIDITSTERTDVENTICSSSSSRVAMSYN 144  
b 61 TITLMLRRNLNLYVMGYSDFPENGKCRHYFNIDITSTERTDVENTICSSSSSRVAMSYN 119  
Y 145 SLYPTMEKKAENSRNOVGLGIGLSSDICKISGVDSFPVKTEAFELLVAIQWSEAAHF 204  
b 120 GLYPTLEKKAGVTSRNEVGLGIGLSSDICKISGVDSFPVKTEAFELLVAIQWSEAAHF 179  
Y 205 KYIENQVKTNFRAYDPDKVNLNLEKWKISIAHNKNGALPKPPLVLDAGTKWIVL 264  
b 180 KYIENQVKTNFRAYDPDKVNLNLEKWKISIAHNKNGALPKPPLVLDAGTKWIVL 239  
Y 265 RVDEINRVALKYNGTCQT 286  
b 240 RVDEIKPDVGLNLYNGTCQT 261

RESULT 12  
US-08-646-360-9  
Sequence 9, Application US/08646360  
Patent No. 5837491  
GENERAL INFORMATION:  
APPLICANT: Better, Marc D.  
APPLICANT: Carroll, Stephen F.  
APPLICANT: Studnika, Gary M.  
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
NUMBER OF SEQUENCES: 173  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
STREET: 500 West Madison Street, 34th floor  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60661  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/646,360  
FILING DATE: 13-MAY-1996  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/05348

FILING DATE: 12-MAY-1994  
APPLICATION NUMBER: US 08/054,691  
FILING DATE: 12-MAY-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/988,430  
FILING DATE: 09-DEC-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: McNicholas, Janet M.  
REGISTRATION NUMBER: 32,918  
REFERENCE/DOCKET NUMBER: 200-70.P4  
TELEPHONE: 312/707-8889  
TELEFAX: 312/707-9155  
TELEX: 650 388-1248  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 261 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-646-360-9

Query Match 67.8%; Score 1096.5; DB 2; Length 261;  
Best Local Similarity 82.8%; Pred. No. 1.5e-101;  
Matches 217; Conservative 14; Mismatches 30; Indels 1; Gaps 1;  
Y 25 INTITFDAGNATINKYATPMESLRNOAKDPKLCYGPMLPDTNSTPKYLLVKLOGANLK 84  
b 1 INTITFDAGNATINKYATPMESLRNEAKDPSLKCYGIPMLPNTNSTIKYLLVKLOGASLK 60  
Y 85 TITLMLRRNLNLYVMGYSDFPENGKCRHYFNIDITSTERTDVENTICSSSSSRVAMSYN 144  
b 61 TITLMLRRNLNLYVMGYSDFPENGKCRHYFNIDITSTERTDVENTICSSSSSRVAMSYN 119  
Y 145 SLYPTMEKKAENSRNOVGLGIGLSSDICKISGVDSFPVKTEAFELLVAIQWSEAAHF 204  
b 120 GLYPTLEKKAGVTSRNEVGLGIGLSSDICKISGVDSFPVKTEAFELLVAIQWSEAAHF 179  
Y 205 KYIENQVKTNFRAYDPDKVNLNLEKWKISIAHNKNGALPKPPLVLDAGTKWIVL 264  
b 180 KYIENQVKTNFRAYDPDKVNLNLEKWKISIAHNKNGALPKPPLVLDAGTKWIVL 239  
Y 265 RVDEINRVALKYNGTCQT 286  
b 240 RVDEIKPDVGLNLYNGTCQT 261

RESULT 13  
US-08-839-765-9  
Sequence 9, Application US/08839765  
Patent No. 6146631  
GENERAL INFORMATION:  
APPLICANT: Better, Marc D.  
APPLICANT: Carroll, Stephen F.  
APPLICANT: Studnika, Gary M.  
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
NUMBER OF SEQUENCES: 169  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
STREET: 500 West Madison Street, 34th floor  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60661  
COMPUTER READABLE FORM:

TELEX: 650 388-1248  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 263 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-488-113B-9

Query Match 67.8% Score 1096.5 DB 1: Length 261:  
Best Local Similarity 82.8% Pred. No. 1.5e-101:  
Matches 217; Conservative 14; Mismatches 30; Indels 1; Gaps 1:

25 INTITDAGNATINKYATFMSLRNKAQKPKLKYGIPMLPDTNSTPKYLIVKLOCANLK 84  
1 INTITDAGNATINKYATFMSLRNKAQKPKLKYGIPMLPDTNSTPKYLIVKLOCANLK 60

85 TITMLRRNLNLYVMGYSDPFNGNKCRIYHIFNDITSTERTDVENTLCCSSSSSRVAMSYN 144  
61 TITMLRRNLNLYVMGYSDPYD-NKCRYHIFNDIKCTEYSDVENTLCPSSNPRVAKPINYN 119

145 SLYPTMEKKAENSRNQVQLGTOILSSDICKISGVDSFPVKTEAFELLVAIQWSEAAAF 204  
120 GLYPTLEKKAGVTSRNEVQLGTOILSSDICKISGVDSFPVKTEAFELLVAIQWSEAAAF 179

205 KYTENOVKTNFNRFPDVKVINLEEKWKISEAIIHNAKNGALPKPLELVDAGTKWIVL 264  
180 KYTENOVKTNFNRDFSPNDKVLDEENWCKISTAIHNSKNGALPKPLELKNADGTKWIVL 239

265 RVDEINRDVALLKYVNGTCOTT 286  
240 RVDEIKPDVGLLYVNGTCOAT 261

RESULT 10  
US-08-488-113B-9  
Sequence 9, Application US/08488113B  
Patent No. 5744580  
GENERAL INFORMATION:  
APPLICANT: Better, Marc D.  
APPLICANT: Carroll, Stephen F.  
APPLICANT: Studnika, Gary M.  
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
NUMBER OF SEQUENCES: 169  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
STREET: 500 West Madison Street, 34th floor  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60661  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/425,336  
FILING DATE: 18-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/064,691  
FILING DATE: 12-MAY-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/988,430  
FILING DATE: 09-DEC-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: McNicholas, Janet M.  
REGISTRATION NUMBER: 32,918  
REFERENCE/DOCKET NUMBER: 11022US07/200-70.P3.C3A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/707-8889  
TELEFAX: 312/707-9155

TELEX: 650 388-1248  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 263 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-488-113B-9

Query Match 67.8% Score 1096.5 DB 1: Length 261:  
Best Local Similarity 82.8% Pred. No. 1.5e-101:  
Matches 217; Conservative 14; Mismatches 30; Indels 1; Gaps 1:

25 INTITDAGNATINKYATFMSLRNKAQKPKLKYGIPMLPDTNSTPKYLIVKLOCANLK 84  
1 INTITDAGNATINKYATFMSLRNKAQKPKLKYGIPMLPDTNSTPKYLIVKLOCANLK 60

85 TITMLRRNLNLYVMGYSDPFNGNKCRIYHIFNDITSTERTDVENTLCCSSSSSRVAMSYN 144  
61 TITMLRRNLNLYVMGYSDPYD-NKCRYHIFNDIKCTEYSDVENTLCPSSNPRVAKPINYN 119

145 SLYPTMEKKAENSRNQVQLGTOILSSDICKISGVDSFPVKTEAFELLVAIQWSEAAAF 204  
120 GLYPTLEKKAGVTSRNEVQLGTOILSSDICKISGVDSFPVKTEAFELLVAIQWSEAAAF 179

205 KYTENOVKTNFNRFPDVKVINLEEKWKISEAIIHNAKNGALPKPLELVDAGTKWIVL 264  
180 KYTENOVKTNFNRDFSPNDKVLDEENWCKISTAIHNSKNGALPKPLELKNADGTKWIVL 239

265 RVDEINRDVALLKYVNGTCOTT 286  
240 RVDEIKPDVGLLYVNGTCOAT 261

RESULT 11  
US-08-477-484B-9  
Sequence 9, Application US/08477484B  
Patent No. 5756699  
GENERAL INFORMATION:  
APPLICANT: Better, Marc D.  
APPLICANT: Carroll, Stephen F.  
APPLICANT: Studnika, Gary M.  
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
NUMBER OF SEQUENCES: 169  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
STREET: 500 West Madison Street, 34th floor  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60661  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/425,336  
FILING DATE: 18-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/064,691  
FILING DATE: 12-MAY-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/988,430  
FILING DATE: 09-DEC-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992

SRNT

Printed 11/17/2004

Query Match 75.9% Score 1227.5 DB 4: Length 313:  
Best Local Similarity 75.6% Pred. NO. 1.6e-114: Indels 3: Gaps 2:  
Matches 238: Conservative 33: Mismatches 41: Indels 3: Gaps 2:  
1 MKVMLVVVTVLAWLZAAPTSCAINTITFDAGNATINKYATFMSLRNOAKDPKLCYCG 60  
1 MKSMLVVTVSI--WLIAPTSTWAVNTIYVVGSTISKYATFMRNLNRNEAKDPSLCYCG 58  
61 IMLPDTNSTPKYLLVKKLOGANIKTTITMLRRNLNVYMGYSDFPNKCRHYHFNIDTST 120  
59 IMLPNTNPNKHLVVELOGSNKKTTITMLRRNLNVYMGYSDFPNKCRHYHFNIDTST 118  
121 ERTDVENTICSSSSSRVANSINYSLYPTMEKKAENSRNOVOLGQIILSSDICKISGVND 180  
119 ERQDVEITLCPNANSRVSKNIINFSRYPTLESKAGVKSQVOLGQIILSDNIGKISGVN 178  
181 SPVKTAEAFELLVAIQVSEAAARFKYIENQVKTNFNRAFYDPKVINLEEKMGKISEATH 240  
179 SITEKTEAEFELLVAIQVSEAAARFKYIENQVKTNFNRAFYDPKVINLEEKMGKISEATH 238  
241 NAKNGALPKPLELDVAKGTWVLRVDEINROVALLKYVNGTCQTTY-ONAMFSQVLIIST 299  
239 DAKNGVLPKPLELDVAGAKWVLRVDEIKPDVALLNYYVGGSCQTTYONAMFPOLIMST 298  
300 YNYMSNLGLDFEGF 314  
299 YNYMVNGLDFEGF 313

RESULT 8  
US-08-378-761A-79  
Sequence 79, Application US/08378761A  
Patent No. 5635384  
GENERAL INFORMATION:  
APPLICANT: WALSH, TERENCE A  
APPLICANT: HEY, TIMOTHY D  
APPLICANT: MORGAN, ALICE ER  
TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE  
TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF  
NUMBER OF SEQUENCES: 81  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ANDREA T. BORUCKI  
STREET: 9330 ZIONSVILLE ROAD  
CITY: INDIANAPOLIS  
STATE: IN  
COUNTRY: US  
ZIP: 46268  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/378,761A  
FILING DATE: 26-JAN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: BORUCKI, ANDREA T  
REGISTRATION NUMBER: 33651  
REFERENCE/DOCKET NUMBER: 38272B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (317) 337-4846  
INFORMATION FOR SEQ ID NO: 79:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 261 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-378-761A-79

Query Match 67.8% Score 1096.5 DB 1: Length 261:  
Best Local Similarity 82.8% Pred. NO. 1.5e-101:  
Matches 217: Conservative 14: Mismatches 30: Indels 1: Gaps 1:  
25 INTITFDAGNATINKYATFMSLRNOAKDPKLCYCGIPLPDTNSTPKYLLVKKLOGANIK 84  
1 INTITFDAGNATINKYATFMSLRNEAKDPSLCYCGIPLPDTNSTIKYLLVKKLOGANIK 60  
85 TITMLRRNLNVYMGYSDFPNKCRHYHFNIDTSTERTDVENTICSSSSSRVANSINYN 144  
61 TITMLRRNLNVYMGYSDFPNKCRHYHFNIDTSTERTDVENTICSSSSSRVANSINYN 119  
145 SLVPTMEKKAENSRNOVOLGQIILSSDICKISGVDSFPVKTEAFELLVAIQVSEAAARF 204  
120 GLYPTLEKACVTSRNEVOLGQIILSSDICKISGVDSFPVKTEAFELLVAIQVSEAAARF 179  
205 KYIENQVKTNFNRAFYDPKVINLEEKMGKISEATHNKGALPKPLELDVAKGTWVLRV 264  
180 KYIENQVKTNFNRAFYDPKVINLEEKMGKISEATHNKGALPKPLELDVAKGTWVLRV 239  
265 RVDEINROVALLKYVNGTCQTTY 286  
240 RVDEJKPDVGLLVNNGTCQAT 261

RESULT 9  
US-08-485-286-79  
Sequence 79, Application US/08485286  
Patent No. 5646026  
Patent No. 5646026 5646119  
GENERAL INFORMATION:  
APPLICANT: WALSH, TERENCE A  
APPLICANT: HEY, TIMOTHY D  
APPLICANT: MORGAN, ALICE ER  
TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE  
TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF  
NUMBER OF SEQUENCES: 81  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ANDREA T. BORUCKI  
STREET: 9330 ZIONSVILLE ROAD  
CITY: INDIANAPOLIS  
STATE: IN  
COUNTRY: US  
ZIP: 46268  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/485,286  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/378761  
FILING DATE: 26-JAN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: BORUCKI, ANDREA T  
REGISTRATION NUMBER: 33651  
REFERENCE/DOCKET NUMBER: 38272B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (317) 337-4846  
INFORMATION FOR SEQ ID NO: 79:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 261 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-485-286-79

Query Match 67.8% Score 1096.5 DB 1: Length 261:

Matches 239; Conservative 33; Mismatches 40; Indels 3; Gaps 2;

1 MKVMLVVVVVTLIAWLIAPTSTCAINTITFDAGNATINKYATFMSLRNQAKDPKLCYG 60  
1 MKSMLVVTISI--WLLAPTSTWAVNTIIVYVGGSTTSKYATFLNDRNEAKDPSLKCYG 58  
61 IMLPDNTNPKYLVKLOGANKTITLMLRNNLYVMGYSDPFENKCRHYHENDITST 120  
59 IMLPDNTNPKYLVKLOGANKTITLMLRNNLYVMGYSDPFENKCRHYHENDISCT 118  
121 BRTDVENTLCCSSSSSRVAMSINYSIYPTMEKKAENSRNQVQLGIQLSSDGIKISVD 180  
119 ERQDVETTLCPNANRSVSKNINFDSPRYPLESKAGVRSQVQLGIQLSSDGIKISVD 178  
181 SPPVKTEAFFLLVAIOMYSEARFKYIENOVKTENRAFYPDPKVINLEEKWKI SEATH 240  
179 SFTKTEAFLVLAOMYSEARFKYIENOVKTENRAFNPKNPKVNLQETWCKISTAIH 238  
241 NAKNGALPKPLEVDKAGTKWIVLRVDEINRDVALLKYVNGTCOTTY-ONAMFSQVIIST 299  
239 DAKNGVLPKPLEVDKAGTKWIVLRVDEIKPDVALLNYVGGSCQTTYNQAMFPOLINST 298  
300 YNYMNLGDLFEFG 314  
299 YNYMNLGDLFEFG 313

RESULT 6  
US-08-501-253A-2  
Sequence 2, Application US/08501253A  
Patent No. 6146628  
GENERAL INFORMATION:  
APPLICANT: Uckun, Fatih  
TITLE OF INVENTION: Bioreproductive Agents Comprising Recombinant PAP and PAP Mutants  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Merchant & Gould  
STREET: 90 South 7th Street, 3100 No. 6146628west Center  
CITY: Minneapolis  
STATE: MN  
COUNTRY: US  
ZIP: 55402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/501,253A  
FILING DATE: 11-JUL-1995  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Kettleberger, Denise M.  
REGISTRATION NUMBER: 33,924  
REFERENCE/DOCKET NUMBER: 600.323US01  
TELEPHONE: 612-332-5300  
TELEFAX: 612-332-9081  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 313 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-501-253A-2

Query Match 76.2% Score 1232.5; DB 4: Length 313;  
Identical Similarity 75.6%; Pred. No. 5e-115;  
Matches 238; Conservative 34; Mismatches 40; Indels 3; Gaps 2;

1 MKVMLVVVVVTLIAWLIAPTSTCAINTITFDAGNATINKYATFMSLRNQAKDPKLCYG 60

121 BRTDVENTLCCSSSSSRVAMSINYSIYPTMEKKAENSRNQVQLGIQLSSDGIKISVD 180  
119 ERQDVETTLCPNANRSVSKNINFDSPRYPLESKAGVRSQVQLGIQLSSDGIKISVD 178  
181 SPPVKTEAFFLLVAIOMYSEARFKYIENOVKTENRAFYPDPKVINLEEKWKI SEATH 240  
179 SFTKTEAFLVLAOMYSEARFKYIENOVKTENRAFNPKNPKVNLQETWCKISTAIH 238  
241 NAKNGALPKPLEVDKAGTKWIVLRVDEINRDVALLKYVNGTCOTTY-ONAMFSQVIIST 299  
239 DAKNGVLPKPLEVDKAGTKWIVLRVDEIKPDVALLNYVGGSCQTTYNQAMFPOLINST 298  
300 YNYMNLGDLFEFG 314  
299 YNYMNLGDLFEFG 313

RESULT 7  
US-09-005-273-4  
Sequence 4, Application US/09005273  
Patent No. 6137030  
GENERAL INFORMATION:  
APPLICANT: Tumer, Nilgun E.  
TITLE OF INVENTION: PAP MUTANTS THAT EXHIBIT ANTI-VIRAL AND/OR ANTI-FUNGAL ACTIVITY IN PLANTS  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LERNER, DAVID, LITTENBERG, KRUMHOLZ 6  
STREET: 600 South, Avenue West  
CITY: Westfield  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07090  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/005,273  
FILING DATE: 09-JAN-1998  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/500,611  
FILING DATE: 11-JUL-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/500,694  
FILING DATE: 11-JUL-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/11546  
FILING DATE: 11-JUL-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Foley, Shawn P.  
REGISTRATION NUMBER: 33,071  
REFERENCE/DOCKET NUMBER: OCIRS 3.3-034  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 908-654-5000  
TELEFAX: 908-654-7866  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 313 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-005-273-4



1 MKSMLVVTISI--WLLAPTSTWAVNTIYVGVSTTISKYATFLNDRNEAKDPSLKCYG 58  
61 IMLPDTNSTPKYLVKLGANKLTIITMLRRNNLVYMGYSOPFNGNCKRYHIFNDITST 120  
59 IMLPNTNPKYLVVLOGSNKKTTITMLRRNNLVYMGYSOPFETNCKRYHIFNDISGT 118  
121 ERTDVENTLCSSSSSRVAMSINYSLYPTMEKKAENSHNOVOLGQILSSDGTGKISGVD 180  
119 ERDQVETTLCPNANSRVKSNINFDSPRYPTLESKAGVKSRSQVOLGQILDSNIGKISGVM 178  
181 SFVKTEAEFLVAIQMVSEAAEFKYIENQVKTENRAFYPPDKVINLEEKWKISEAIIH 240  
179 SPTEKTEAEFLVAIQMVSEAAEFKYIENQVKTENRAFPNPNPKVLNLQETWCKISTAIH 238  
241 NAKNGALPKPLELDVAKGTWIVLRVDEINRDVALLKYVNGTCQTTY-ONAMFSQVIST 299  
239 DAKNGVLPKPLELDVDSAGAKWIVLRVDEIKPDVALLNYVGGSCQTTYNONAMFPQLIMST 298  
300 YNYMSNLGDLFEFG 314  
299 YNYMVNUGDLFEFG 313

RESULT 4

JS-09-005-273-2  
: Sequence 2, Application US/09005273  
: Patent No. 613/030  
: GENERAL INFORMATION:  
: APPLICANT: Tumer, Nilgun E.  
: TITLE OF INVENTION: PAP MUTANTS THAT EXHIBIT ANTI-VIRAL  
: AND/OR ANTI-FUNGAL ACTIVITY IN PLANTS  
: NUMBER OF SEQUENCES: 4  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: LERNER, DAVID, LITTENBERG, KRUMHOLZ 6  
: ADDRESSEE: MENTILIK  
: STREET: 600 South, Avenue West  
: CITY: Westfield  
: STATE: New Jersey  
: COUNTRY: USA  
: ZIP: 07090

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/005,273  
FILING DATE: 09-JAN-1998  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/500,611  
FILING DATE: 11-JUL-1995.

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/500,694  
FILING DATE: 11-JUL-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/11546  
FILING DATE: 11-JUL-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Foley, Shawn P.  
REGISTRATION NUMBER: 33,071  
REFERENCE/DOCKET NUMBER: OCIRS 3.3-014  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 908-654-5000  
TELEFAX: 908-654-7866  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 313 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-005-273-2

Query Match 76.4% Score 1235.5; DB 4; Length 313;  
Best Local Similarity 75.9% Pred. No. 2.5e-115;  
Matches 239; Conservative 33; Mismatches 40; Indels 2;  
Gaps 2;  
QY 1 MKVMLVVVTLIAWLIAPTSTCAINTITFDAGNATINKYATFENESLRNOAKDPSLKCYG 60  
DB 1 MKSMLVVTISI--WLLAPTSTWAVNTIYVGVSTTISKYATFLNDRNEAKDPSLKCYG 58  
QY 61 IMLPDTNSTPKYLVKLGANKLTIITMLRRNNLVYMGYSOPFNGNCKRYHIFNDITST 120  
DB 59 IMLPNTNPKYLVVLOGSNKKTTITMLRRNNLVYMGYSOPFETNCKRYHIFNDISGT 118  
QY 121 ERTDVENTLCSSSSSRVAMSINYSLYPTMEKKAENSHNOVOLGQILSSDGTGKISGVD 180  
DB 119 ERDQVETTLCPNANSRVKSNINFDSPRYPTLESKAGVKSRSQVOLGQILDSNIGKISGVM 178  
QY 181 SFVKTEAEFLVAIQMVSEAAEFKYIENQVKTENRAFYPPDKVINLEEKWKISEAIIH 240  
DB 179 SPTEKTEAEFLVAIQMVSEAAEFKYIENQVKTENRAFPNPNPKVLNLQETWCKISTAIH 238  
QY 241 NAKNGALPKPLELDVAKGTWIVLRVDEINRDVALLKYVNGTCQTTY-ONAMFSQVIST 299  
DB 239 DAKNGVLPKPLELDVDSAGAKWIVLRVDEIKPDVALLNYVGGSCQTTYNONAMFPQLIMST 298  
QY 300 YNYMSNLGDLFEFG 314  
DB 299 YNYMVNUGDLFEFG 313

RESULT 5

PCT-US96-11546-2  
: Sequence 2, Application PC/TUS9611546  
: GENERAL INFORMATION:  
: APPLICANT: Tumer, Nilgun E.  
: TITLE OF INVENTION: DNAs Encoding Poxeeweed Antiviral Protein  
: NUMBER OF SEQUENCES: 2  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Lerner, David, Littenberg, Krumholz & Mentlik  
: STREET: 600 South Avenue West  
: CITY: Westfield  
: STATE: NJ  
: COUNTRY: USA  
: ZIP: 07090-1497

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/11546  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/500,694  
FILING DATE: 11-JUL-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Foley, Shawn P.  
REFERENCE/DOCKET NUMBER: OCIRS 3.0-039  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 908-654-5000  
TELEFAX: 908-654-7866  
TELEX: 139-125  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 313 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US96-11546-2

Query Match 76.4% Score 1235.5; DB 5;  
Best Local Similarity 75.9% Pred. No. 2.5e-115;

Length 313;

TISSUE TYPE: Leaf

us-09-978-274a-2

Query Match 76.4%; Score 1235.5; DB 1; Length 313;

Best Local Similarity 75.9%; Pred. No. 2.5e-115;

Matches 239; Conservative 33; Mismatches 40; Indels 3; Gaps 2;

Qy 1 MKVMLVVVTLIAWLIAPTSTCAINTITFDAGNATINKYATFESLRNQAOKPKLCYCG 60  
Db 1 MKSMLVVTISI--WLILAPTSTWAVNTIYNVGSSTISKYATFELNDRNEAKDPSLCYCG 58  
Qy 61 IPMLPDTNSTPKYLLVLOGANLKTITLMLRRNNLYVMGYSDPENGKCRHYHNDITST 120  
Db 59 IPMLPDTNSTPKYLLVLOGANLKTITLMLRRNNLYVMGYSDPENGKCRHYHNDITST 118  
Qy 121 ERTDVENTLCSSSSSRVAMSNLYPTMEKKAENVSRNOVOLGQILSSDICKISGVD 180  
Db 119 ERODVEITLCPNANRSVSKNINFDOSRYPTLESKAGVKSRSQVQLGQILDSNICKISGVM 178  
Qy 181 SPVKTEAFFLLVAIQNVSEAAAEKYIENQVKTNFRNRAFYDPDPKVINLEEKWKGKISEAIH 240  
Db 179 SFTEKTEAEFLVAIQNVSEAAAEKYIENQVKTNFRNRAFYDPDPKVINLEEKWKGKISEAIH 238  
Qy 241 NAKNGALPKLELVDAKCTKWILRVDEINRDVALLKYVNGTCQTTY--QNAMFQSVIIST 299  
Db 239 DAKNGVLPKPLELVDAKCAKWIIVLVDKLPDVALNLYVGSQCTTYNQNAMFPQLIMST 298  
Qy 300 YYNYSNLGDLPEGF 314  
Db 299 YYNVMNLGDLPEGF 313

RESULT 2

Sequence 2, Application US/08500611

Patent No. 5756322

GENERAL INFORMATION:

APPLICANT: Turner, Nilgun E.  
TITLE OF INVENTION: Pokeweed Antiviral Protein Mutants  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lerner, David, Littenberg, Krumholz & Mentlik  
STREET: 600 South Avenue West  
CITY: Westfield  
STATE: NJ  
COUNTRY: USA  
ZIP: 07090-1497

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
FILING DATE: 11-JUL-1995  
APPLICATION NUMBER: US/08/500.611  
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Foley, Shawn P.  
REFERENCE/DOCKET NUMBER: OCIRS 3.0-034  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 908-654-5000  
TELEFAX: 908-654-7866  
TELEX: 139-125

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:  
LENGTH: 313 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein

us-09-500-611-2

Query Match 76.4%; Score 1235.5; DB 1; Length 313;

Best Local Similarity 75.9%; Pred. No. 2.5e-115;

Matches 239; Conservative 33; Mismatches 40; Indels 3; Gaps 2;

Qy 1 MKVMLVVVTLIAWLIAPTSTCAINTITFDAGNATINKYATFESLRNQAOKPKLCYCG 60  
Db 1 MKSMLVVTISI--WLILAPTSTWAVNTIYNVGSSTISKYATFELNDRNEAKDPSLCYCG 58  
Qy 61 IPMLPDTNSTPKYLLVLOGANLKTITLMLRRNNLYVMGYSDPENGKCRHYHNDITST 120  
Db 59 IPMLPDTNSTPKYLLVLOGANLKTITLMLRRNNLYVMGYSDPENGKCRHYHNDITST 118  
Qy 121 ERTDVENTLCSSSSSRVAMSNLYPTMEKKAENVSRNOVOLGQILSSDICKISGVD 180  
Db 119 ERODVEITLCPNANRSVSKNINFDOSRYPTLESKAGVKSRSQVQLGQILDSNICKISGVM 178  
Qy 181 SPVKTEAFFLLVAIQNVSEAAAEKYIENQVKTNFRNRAFYDPDPKVINLEEKWKGKISEAIH 240  
Db 179 SFTEKTEAEFLVAIQNVSEAAAEKYIENQVKTNFRNRAFYDPDPKVINLEEKWKGKISEAIH 238  
Qy 241 NAKNGALPKLELVDAKCTKWILRVDEINRDVALLKYVNGTCQTTY--QNAMFQSVIIST 299  
Db 239 DAKNGVLPKPLELVDAKCAKWIIVLVDKLPDVALNLYVGSQCTTYNQNAMFPQLIMST 298  
Qy 300 YYNYSNLGDLPEGF 314  
Db 299 YYNVMNLGDLPEGF 313

RESULT 3

US-08-500-694-2

Sequence 2, Application US/08500694

Patent No. 5880329

GENERAL INFORMATION:

APPLICANT: Turner, Nilgun E.  
TITLE OF INVENTION: DNAS Encoding Pokeweed Antiviral Protein  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lerner, David, Littenberg, Krumholz & Mentlik  
STREET: 600 South Avenue West  
CITY: Westfield  
STATE: NJ  
COUNTRY: USA  
ZIP: 07090-1497

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
FILING DATE: 11-JUL-1995  
APPLICATION NUMBER: US/08/500.694  
CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: Foley, Shawn P.  
REFERENCE/DOCKET NUMBER: OCIRS 3.0-039  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 908-654-5000  
TELEFAX: 908-654-7866  
TELEX: 139-125

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:  
LENGTH: 313 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein

US-08-500-694-2

Query Match 76.4%; Score 1235.5; DB 2; Length 313;

Best Local Similarity 75.9%; Pred. No. 2.5e-115;

Matches 239; Conservative 33; Mismatches 40; Indels 3; Gaps 2;

Qy 1 MKVMLVVVTLIAWLIAPTSTCAINTITFDAGNATINKYATFESLRNQAOKPKLCYCG 60  
Db 1 MKSMLVVTISI--WLILAPTSTWAVNTIYNVGSSTISKYATFELNDRNEAKDPSLCYCG 58



RESULT 15  
 AAX81496/c  
 ID AAX81496 standard: DNA: 1956 BP.  
 XX  
 AC AAX81496;  
 XX  
 DT 24-AUG-1999 (first entry)  
 XX  
 DE DNA sequence of cry3Bb.11058 gene.  
 XX  
 DE Cry3Bb; mutant; insecticidal activity; insecticidal specificity;  
 KW coleoptera; southern corn rootworm; western corn root worm;  
 KW Diabrotica undecimpunctata howardi Barber; transgenic plant;  
 KW Diabrotica virgifera virgifera LeConte; insecticide resistance; ss.  
 XX  
 OS Synthetic.  
 OS Bacillus thuringiensis.  
 XX  
 PN W09931248-A1.  
 XX  
 PD 24-JUN-1999.  
 XX  
 PF 17-DEC-1998; 98WO-US26852.  
 XX  
 PR 18-DEC-1997; 97US-0996441.  
 PR 18-DEC-1997; 97US-0993170.  
 PR 18-DEC-1997; 97US-0993722.  
 PR 18-DEC-1997; 97US-0993775.  
 XX  
 PA (ECOC-) ECOGEN INC.  
 PA (MONS ) MONSANTO CO.  
 XX  
 PJ Brussock SM, Bryson JW, English L, Kulesza CA, Malvar TM;  
 PI Romano C, Slatin SL, Von Tersch MA, Walters FS;  
 XX  
 DR WPI: 1999-395184/33.  
 XX  
 PT Insecticidal Bacillus thuringiensis proteins  
 XX  
 Claim 52: Page 426-429; 512pp; English.  
 XX  
 PS AAX01468-x81502, AAX81529 and AAX81532 encode new Bacillus thuringiensis  
 CC Cry3Bb mutant proteins. The specification also describes methods of  
 CC altering Bacillus thuringiensis Cry3Bb. The B. thuringiensis Cry3Bb  
 CC polypeptide was modified to have improved insecticidal activity or  
 CC enhanced insecticidal specificity against a target insect. The  
 CC modification comprises at least one amino acid substitution, addition,  
 CC or deletion in the primary sequence of the native or unmodified Cry3Bb  
 CC polypeptide, wherein the substitution or deletion occurs at a position  
 CC corresponding to from about amino acids 1-365 of the unmodified  
 CC polypeptide sequence (AAX81528 and AAX81531 encode the wild type Cry3Bb  
 CC protein). The polypeptide can be used to kill coleopteran pests,  
 CC especially by application to the environment. It is especially  
 CC useful against southern corn rootworm and western corn root worm,  
 CC (Diabrotica undecimpunctata howardi Barber, and Diabrotica virgifera  
 CC virgifera LeConte respectively). The mutant cry3Bb polynucleotides  
 CC can also be used to produce transgenic plants with increased  
 CC insecticide resistance.  
 XX  
 SQ Sequence 1956 BP: 683 A: 315 C: 344 G: 614 T: 0 other;  
 Query Match 2.1% Score 20; DB 20; Length 1956;  
 Best Local Similarity 100.0%; Pred. No. 8;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 58 ACTTCAACTGTGCCATAAA 77  
 DB 344 ACTTCAACTGTGCCATAAA 325  
 Search completed: June 2, 2003, 04:22:00  
 Job time : 292 secs

Matches 22: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 844 ACTGTCAGACAACTTACCATAA 865  
 Db 1974 ACTGTCAGACAACTTACCATAA 1995

RESULT 13  
 AAX81493/C  
 ID AAX81493 standard: DNA: 1956 BP.  
 XX AAX81493;  
 XX  
 DT 24-AUG-1999 (first entry)  
 XX  
 DE DNA sequence of cry3Bb.11048 gene.  
 XX  
 KW Cry3Bb: mutant; insecticidal activity; insecticidal specificity;  
 KW coleoptera; southern corn rootworm; western corn root worm;  
 KW Diabrotica undecimpunctata howardi Barber; transgenic plant;  
 KW Diabrotica virgifera virgifera LeConte; insecticide resistance; ss.  
 XX  
 OS Synthetic.  
 OS Bacillus thuringiensis.  
 XX  
 PN WO9931248-A1.  
 XX  
 PD 24-JUN-1999.  
 XX  
 PF 17-DEC-1998; 98MO-US26852.  
 XX  
 PR 18-DEC-1997; 97US-0996441.  
 PR 18-DEC-1997; 97US-0993170.  
 PR 18-DEC-1997; 97US-0993722.  
 PR 18-DEC-1997; 97US-0993775.  
 XX  
 PA (ECOG-) ECOGEN INC.  
 PA (MONS ) MONSANTO CO.  
 XX  
 PI Brussock SM, Bryson JW, English L, Kulesza CA, Malvar TM;  
 PI Romano C, Slatin SL, Von Tersch MA, Walters FS;  
 XX  
 DR WPI; 1999-395184/33.  
 XX  
 PT Insecticidal Bacillus thuringiensis proteins  
 XX  
 PS Claim 52: Page 409-412; 512pp; English.  
 XX  
 CC AAX81468-X81502, AAX81529 and AAX81532 encode new Bacillus thuringiensis  
 CC Cry3Bb mutant proteins. The specification also describes methods of  
 CC altering Bacillus thuringiensis Cry3Bb. The B. thuringiensis Cry3Bb  
 CC polypeptide was modified to have improved insecticidal activity or  
 CC enhanced insecticidal specificity against a target insect. The  
 CC modification comprises at least one amino acid substitution, addition,  
 CC or deletion in the primary sequence of the native or unmodified Cry3Bb  
 CC polypeptide, wherein the substitution or deletion occurs at a position  
 CC corresponding to from about amino acids 1-365 of the unmodified  
 CC polypeptide sequence (AAX81528 and AAX81531 encode the wild type Cry3Bb  
 CC protein). The polypeptide can be used to kill coleopteran pests,  
 CC especially by application to the environment. It is especially  
 CC useful against southern corn rootworm and western corn root worm,  
 CC (Diabrotica undecimpunctata howardi Barber, and Diabrotica virgifera  
 CC virgifera LeConte respectively). The mutant cry3Bb polynucleotides  
 CC can also be used to produce transgenic plants with increased  
 CC insecticide resistance.  
 XX  
 SQ Sequence 1956 BP: 687 A; 312 C; 346 G; 611 T; 0 other;

Query Match 2.1%; Score 20; DR 20; Length 1956;  
 Best Local Similarity 100.0%; Pred. No. 8;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 58 ACTTCAACTTGTGCCATAAA 77  
 Db 344 ACTTCAACTTGTGCCATAAA 325

Db 344 ACTTCAACTTGTGCCATAAA 325  
 RESULT 14  
 AAX81495/C  
 ID AAX81495 standard: DNA: 1956 BP.  
 XX AAX81495;  
 XX  
 DT 24-AUG-1999 (first entry)  
 XX  
 DE DNA sequence of cry3Bb.11057 gene.  
 XX  
 KW Cry3Bb: mutant; insecticidal activity; insecticidal specificity;  
 KW coleoptera; southern corn rootworm; western corn root worm;  
 KW Diabrotica undecimpunctata howardi Barber; transgenic plant;  
 KW Diabrotica virgifera virgifera LeConte; insecticide resistance; ss.  
 XX  
 OS Synthetic.  
 OS Bacillus thuringiensis.  
 XX  
 PN WO9931248-A1.  
 XX  
 PD 24-JUN-1999.  
 XX  
 PF 17-DEC-1998; 98MO-US26852.  
 XX  
 PR 18-DEC-1997; 97US-0996441.  
 PR 18-DEC-1997; 97US-0993170.  
 PR 18-DEC-1997; 97US-0993722.  
 PR 18-DEC-1997; 97US-0993775.  
 XX  
 PA (ECOG-) ECOGEN INC.  
 PA (MONS ) MONSANTO CO.  
 XX  
 PI Brussock SM, Bryson JW, English L, Kulesza CA, Malvar TM;  
 PI Romano C, Slatin SL, Von Tersch MA, Walters FS;  
 XX  
 DR WPI; 1999-395184/33.  
 XX  
 PT Insecticidal Bacillus thuringiensis proteins  
 XX  
 PS Claim 52: Page 421-423; 512pp; English.  
 XX  
 CC AAX81468-X81502, AAX81529 and AAX81532 encode new Bacillus thuringiensis  
 CC Cry3Bb mutant proteins. The specification also describes methods of  
 CC altering Bacillus thuringiensis Cry3Bb. The B. thuringiensis Cry3Bb  
 CC polypeptide was modified to have improved insecticidal activity or  
 CC enhanced insecticidal specificity against a target insect. The  
 CC modification comprises at least one amino acid substitution, addition,  
 CC or deletion in the primary sequence of the native or unmodified Cry3Bb  
 CC polypeptide, wherein the substitution or deletion occurs at a position  
 CC corresponding to from about amino acids 1-365 of the unmodified  
 CC polypeptide sequence (AAX81528 and AAX81531 encode the wild type Cry3Bb  
 CC protein). The polypeptide can be used to kill coleopteran pests,  
 CC especially by application to the environment. It is especially  
 CC useful against southern corn rootworm and western corn root worm,  
 CC (Diabrotica undecimpunctata howardi Barber, and Diabrotica virgifera  
 CC virgifera LeConte respectively). The mutant cry3Bb polynucleotides  
 CC can also be used to produce transgenic plants with increased  
 CC insecticide resistance.  
 XX  
 SQ Sequence 1956 BP: 684 A; 316 C; 345 G; 611 T; 0 other;

Query Match 2.1%; Score 20; DR 20; Length 1956;  
 Best Local Similarity 100.0%; Pred. No. 8;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 58 ACTTCAACTTGTGCCATAAA 77  
 Db 344 ACTTCAACTTGTGCCATAAA 325

09978274

Query Match 2.3% Score 22; DB 16; Length 918;  
Best Local Similarity 100.0%; Pred. No. 0.78;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 289 GTGATGGCGATTCTGATCCCT 310  
|||||  
DB 277 GTGATGGCGATTCTGATCCCT 298

RESULT 11  
AAT99557  
ID AAT99557 standard: DNA: 2369 BP.  
XX  
AC AAT99557;  
XX  
DT 08-JUN-1998 (first entry)  
XX  
DE Phytolacca insularis antiviral protein gPIP50 gene.  
XX  
KW Antiviral protein; gPIP50 gene; virucide; transgenic plant;  
KW virus resistance; immunoconjugate; AIDS; cancer; therapy; ss.  
XX  
OS Phytolacca insularis Nakai.

XX Key Location/Qualifiers  
FH CAAT\_signal 983..986  
FT /tag= a  
FT TATA\_signal 1032..1035  
FT /tag= b  
FT TATA\_signal 1090..1093  
FT /tag= c  
FT CDS 1128..2075  
FT /tag= d  
FT /note= "Claim 3"  
FT sig\_peptide 1128..1199  
FT /tag= e  
FT mat\_peptide 1200..2072  
FT /tag= f  
FT polyA\_signal 2142..2146  
FT /tag= g  
FT polyA\_signal 2263..2268  
FT /tag= h

XX EP808902-A2.  
XX  
XX 26-NOV-1997.  
XX  
XX 30-SEP-1996; 96EP-0307159.  
XX  
XX 22-MAY-1996; 96KR-0017404.  
XX  
XX (JINR-) JIN RO LTD.  
XX Chai J, Choi K, Choi Y, Hong E, Jin J, Kim C;  
XX Kim M, Lee J, Lee J, Moon Y, Song S, Yun Y;  
XX  
XX WPI: 1998-001788/01.  
XX P-PSDB: AAW26774.

XX Antiviral proteins of Phytolacca insularis Nakai and their genes  
XX useful in plant antiviral agents and immunoconjugates for the  
XX treatment of AIDS and cancer  
XX  
XX Claim 3: Fig 5: 26pp; English.

XX This polynucleotide comprises the Phytolacca insularis Nakai gene  
XX coding for a 35.7 kDa novel antiviral protein designated gPIP50  
XX (see AAW26774). The gPIP50 gene was isolated from leaf genomic DNA  
XX using the coding region of antiviral protein cPIP as probe.  
XX Another gene (see AAT99556), encoding a 32.8 kDa protein (see AAW26773)  
XX designated gPIP2, has also been isolated from P. insularis Nakai.  
XX Also claimed are vectors encoding these antiviral proteins and host  
XX cells transformed or transfected with these vectors. E. coli

CC xli-blue MRF' gPIP50 (KCCM-10081) host cells are claimed, as is a  
CC process for preparing antiviral protein by cultivating these cells  
CC and purifying the protein from inclusion bodies. The antiviral  
CC proteins and recombinant proteins inhibit protein synthesis. They  
CC can be used as active ingredients of antiviral agents of plant  
CC viruses, and employed in the manufacture of immunoconjugates for  
CC the treatment of AIDS and cancer. The isolated genes can be used  
CC in the breeding of transgenic plants having viral resistance.

XX Sequence 2369 BP: 778 A: 373 C: 450 G: 768 T: 0 other;

Query Match 2.3% Score 22; DB 19; Length 2369;  
Best Local Similarity 100.0%; Pred. No. 0.78;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 844 ACCTGTCAGACAACTTACCAGA 865  
|||||  
DB 1974 ACCTGTCAGACAACTTACCAGA 1995

RESULT 12  
ABA96543  
ID ABA96543 standard: DNA: 2369 BP.  
XX  
AC ABA96543;  
XX  
DT 19-MAR-2002 (first entry)  
XX  
DE Phytolacca insularis antiviral protein (PIP)-encoding DNA.  
XX  
KW Antiviral protein; PIP; cloning vector; ds.  
XX  
XX Phytolacca insularis.  
XX  
XX Key Location/Qualifiers  
FH CDS 1128..2075  
FT /tag= a  
FT /product= "Phytolacca insularis antiviral protein, PIP"  
FT sig\_peptide 1128..1199  
FT /tag= b  
FT mat\_peptide 1200..2072  
FT /tag= c  
FT /product= "Mature PIP"

XX KR98021298-A.  
XX  
XX 25-JUN-1998.  
XX  
XX 16-SEP-1996; 96KR-0040110.  
XX  
XX 16-SEP-1996; 96KR-0040110.  
XX  
XX (JINR-) JIN RO LTD.  
XX Moon YH, Choi JN, Yoon YC, Jin JH, Hong EJ, Lee JH, Chang JI;  
XX Park YC, Choi GH, Kim CH, Song SG, Lee JS, Choi YD;  
XX  
XX WPI: 1999-299028/25.  
XX P-PSDB: AAW53033.  
XX  
XX Cloning vector using gene of Phytolacca insularis antiviral protein  
XX (PIP) -  
XX  
XX Example 1: Fig 1; 9pp; Korean.

XX The invention relates to cloning vectors comprising a gene encoding  
XX the Phytolacca insularis antiviral protein, PIP (ABA96543, AAW53033), or  
XX a fragment thereof. The present sequence represents DNA encoding PIP.  
XX  
XX Sequence 2369 BP: 778 A: 373 C: 450 G: 768 T: 0 other;  
XX  
XX Query Match 2.3% Score 22; DB 20; Length 2369;  
XX Best Local Similarity 100.0%; Pred. No. 0.78;

SRNT

Printed 11/17/2004

DR P-PSDB: AAW26773.  
 XX  
 PT Antiviral proteins of Phytolacca insularis Nakai and their genes  
 PT useful in plant antiviral agents and immunoconjugates for the  
 PT treatment of AIDS and cancer  
 XX  
 PS  
 CC Claim 2: Page 10-12; 26pp; English.  
 CC  
 CC This polynucleotide comprises the coding region of the Phytolacca  
 CC insularis Nakai gene coding for a 32.8 kDa novel antiviral protein  
 CC designated gPIP2 (see AAW26773). The gPIP2 gene was isolated from  
 CC leaf genomic DNA by PCR amplification (see AAT99558-59). Another  
 CC gene (see AAT99557), encoding a 35.7 kDa protein (see AAW26774)  
 CC designated gPIP50, has also been isolated from P. insularis Nakai.  
 CC Also claimed are vectors encoding these antiviral proteins and host  
 CC cells transformed or transfected with these vectors. E. coli  
 CC XL1-Blue MRF<sup>+</sup> gPIP2 (KCM-10080) host cells are claimed, as is a  
 CC process for preparing antiviral protein by cultivating these  
 CC cells and purifying the protein from inclusion bodies. The  
 CC antiviral proteins and recombinant proteins inhibit protein  
 CC synthesis. They can be used as active ingredients of antiviral  
 CC agents of plant viruses, and employed in the manufacture of  
 CC immunoconjugates for the treatment of AIDS and cancer. The  
 CC isolated genes can be used in the breeding of transgenic plants  
 CC having viral resistance.  
 XX  
 SQ Sequence 882 BP; 301 A; 164 C; 174 G; 243 T; 0 other;  
 Query Match 2.4%; Score 23; DB 19; Length 882;  
 Best Local Similarity 100.0%; Pred. No. 0.24;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 289 GTGATGGCGCTATTCTGATCCCTT 311  
 DB 220 GTGATGGCGCTATTCTGATCCCTT 242  
 RESULT 9  
 AAQ64893  
 ID AAQ64893 standard; cDNA; 918 BP.  
 XX  
 AC AAQ64893;  
 XX  
 DT 17-JAN-1995 (first entry)  
 DE Antiviral protein of Phytolacca insularis Nakai.  
 XX  
 KW Antiviral; anti-microbial; bacteriocide; ribosome; inhibition;  
 KW inactivation; Phytolacca insularis Nakai; ss.  
 XX  
 OS Phytolacca insularis Nakai.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..918  
 FT /\*tag= a  
 FT /product= Antiviral protein.  
 XX  
 PN AU648475-B.  
 XX  
 PD 21-APR-1994.  
 XX  
 PF 13-OCT-1993; 93AU-0048972.  
 XX  
 PR 28-AUG-1993; 93KR-0016938.  
 XX  
 PA (JINR-) JIN RO LTD.  
 XX  
 PI Hong-seob J, Kwan-ho L, Kyu-whan C, Man-keun K;  
 PI Young-ho M;  
 XX  
 DR WPI: 1994-167848/21.  
 DR P-PSDB: AAR54839.  
 XX

PT Nucleotide sequence from Phytolacca insularis encoding antiviral  
 PT protein - encodes a ribosome inactivating protein which may be  
 XX used in antimicrobial compns.  
 XX  
 PS Claim 1: Page 13; 22pp; English.  
 XX  
 CC The antiviral protein encoded by this sequence comprises an amino  
 CC acid sequence prevalent in ribosome inactivating proteins (See  
 CC AAK54840). The protein may be used in anti-microbial compositions.  
 XX  
 SQ Sequence 918 BP; 294 A; 177 C; 191 G; 256 T; 0 other;  
 Query Match 2.3%; Score 22; DB 15; Length 918;  
 Best Local Similarity 100.0%; Pred. No. 0.78;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 289 CTCATGCGCTATTCTGATCCCT 310  
 DB 277 CTCATGCGCTATTCTGATCCCT 298  
 RESULT 10  
 AAT04782  
 ID AAT04782 standard; DNA; 918 BP.  
 XX  
 AC AAT04782;  
 XX  
 DT 16-MAY-1996 (first entry)  
 DE DNA pJMC201 fragment encoding P. insularis antiviral protein.  
 XX  
 KW pJMC201: Phytolacca insularis antiviral protein; PIP; Nakai;  
 KW primer: amplify: CMV 35S promoter; transgenic plant; potato;  
 KW Agrobacterium tumefaciens LBA 4404; plant cell transformation; ss.  
 XX  
 OS Phytolacca insularis.  
 XX  
 FH Key Location/Qualifiers  
 FT misc\_difference 913..915  
 FT /\*tag= a  
 FT /codon= seq:gat, aa:ser  
 XX  
 PN AU663031-B.  
 XX  
 PD 21-SEP-1995.  
 XX  
 PF 04-OCT-1994; 94AU-0074404.  
 XX  
 PR 21-JUL-1994; 94KR-0017696.  
 XX  
 PA (JINR-) JIN RO LTD.  
 XX  
 PI Chul-Hwan K, Hong-Seob J, Kyu-Whan C, Man-Keun K;  
 PI Young-Ho M;  
 XX  
 DR WPI: 1995-35885R/47.  
 DR P-PSDB: AAR80106.  
 XX  
 PT Recombinant DNA vector expressing Phytolacca insularis antiviral  
 PT protein - for the production of virus resistant transgenic plants  
 XX  
 PS Example 1: Fig 1; 25pp; English.  
 XX  
 CC This sequence represents a fragment of a recombinant DNA pJMC201 and  
 CC encodes Phytolacca insularis antiviral protein (PIP). This sequence  
 CC was isolated from a cDNA library of Phytolacca insularis Nakai using  
 CC the primer sequences given in AAT04782-84. The amplified sequence was  
 CC used, cloned with the CMV 35S promoter, in the production of a  
 CC transgenic plant, pref. potato, using Agrobacterium tumefaciens LBA  
 CC 4404 for plant cell transformation.  
 XX  
 SQ Sequence 918 BP; 294 A; 177 C; 191 G; 256 T; 0 other;

09978274

XX WPI: 2000-062555/05.  
DR P-PSDB: AAY58025.  
XX  
XX New antiviral DNA useful for generating transgenic plants resistant to  
PT viruses and/or fungi.  
XX  
XX Examples: Page 4-5; 43pp: English.  
XX  
XX This is the Pokeweed antiviral protein (PAP) nucleotide sequence. PAP is  
CC a type I ribosome-inhibiting protein (RIP) found in the cell walls of  
CC Phytolacca americana (pokeweed). It is a single polypeptide chain that  
CC catalytically removes a specific adenine residue from a highly conserved  
CC stem-loop structure in the 28S rRNA of eukaryotic ribosomes. The pokeweed  
CC antiviral protein II (PAP II) protein confers antiviral and or antifungal  
CC activities to plants. A DNA molecule encoding a PAP II protein with an  
CC intact catalytic active site amino acid residue (E172) is useful for  
CC generating transgenic plants. PAP II DNA is useful for generating a  
CC transgenic plants (especially cereal crops) through transforming a  
CC protoplast or introducing the DNA directly into a plant part prior to  
CC regeneration. Cells and seeds which exhibit anti-viral and/or anti-fungal  
CC activity thus have increased resistance to viruses and/or fungi. Viruses  
CC include potato virus X (PVX), cucumber mosaic virus (CMV), and tomato  
CC yellow leaf curl virus, and fungi include Pythium, Phytophthora,  
CC Sclerotinia and Aspergillus. PAP II proteins also confer resistance to  
CC other plants pests including insects, bacteria and nematodes. PAP II DNA  
CC is also useful for identifying a PAP II protein having reduced  
CC cytotoxicity. PAP II has a reduced cytotoxicity compared to PAP, so  
CC unlike PAP transgenic plants which are stunted and sterile, PAP II  
CC transgenic plants have a normal and fertile phenotype.  
XX  
XX Sequence 1379 BP: 488 A; 233 C; 269 G; 389 T; 0 other;

Query Match 3.1%; Score 29; DB 21; Length 1379;  
Best Local Similarity 100.0%; Pred. No. 0.00022;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 915 TAATCTTGGTGATCTATTGAAGGATCT 943  
DB 1136 TAATCTTGGTGATCTATTGAAGGATCT 1164

RESULT 7  
AAC87929  
JD AAC87929 standard: DNA: 1379 BP.  
XX  
XX AAC87929;  
XX  
XX 06-MAR-2001 (first entry)  
XX  
XX P. americana pokeweed antiviral protein encoding DNA SEQ ID NO:1.  
XX  
XX Phytolacca americana; pokeweed; pokeweed antiviral protein; PAP;  
XX cancer; biotherapeutic; fusion protein; immunocombinator; mutant;  
XX cytostatic; anti-HIV; human immunodeficiency virus; AIDS; leukaemia;  
XX lymphoma; brain tumour; neuroblastoma; soft tissue sarcoma;  
XX osteosarcoma; ss.  
XX  
XX Phytolacca americana.  
XX  
XX Key Location/Qualifiers  
XX CDS 225..1166  
XX  
XX /tag= a  
XX /product= "pokeweed antiviral protein (PAP)"  
XX  
XX US6146628-A.  
XX  
XX 14-NOV-2000.  
XX  
XX 11-JUL-1995; 9505-0501253.  
XX  
XX 11-JUL-1995; 9505-0501253.  
XX  
XX

09978274

SRNT

Printed 11/17/2004

PA (MINU) UNIV MINNESOTA & RUTGERS.  
XX (UYNE-) UNIV STATE NEW JERSEY.  
PI Uckun FM, Tumor NE;  
XX WPI: 2001-040422/05.  
DR P-PSDB: AAR36500.  
XX  
XX Immunocombinator useful for treating cancer and acquired  
PT immunodeficiency syndrome, comprises mutant pokeweed anti-viral protein  
PT and a targeting moiety that binds a cell surface receptor.  
XX  
XX Disclosure: Column 47-50; 32pp: English.  
XX  
XX The present invention describes a fusion protein or an immunocombinator  
CC (I), comprising mutant pokeweed Anti-viral protein (PAP) having an  
CC amino acid substitution at residue 75, 97 or 176 of native PAP and a  
CC targeting moiety that binds a cell surface receptor. (I) can have  
CC cytostatic and anti-HIV activities, and is an inhibitor of cellular  
CC RNA or protein synthesis. (I) is useful for treating AIDS and cancers  
CC including leukaemia, lymphoma, a brain tumour, neuroblastoma,  
CC osteosarcoma, soft tissue sarcoma, breast, prostate, ovarian,  
CC testicular, melanoma, lung, or colon cancer. Immunocombinators prepared  
CC using PAP mutants exhibit an improved therapeutic index over  
CC immunocombinators containing either wild-type PAP or variant PAP.  
CC The present sequence encodes the wild-type PAP, which is given in  
CC the exemplification of the present invention.  
XX  
XX Sequence 1379 BP: 488 A; 234 C; 268 G; 389 T; 0 other;

Query Match 3.1%; Score 29; DB 22; Length 1379;  
Best Local Similarity 100.0%; Pred. No. 0.00022;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 915 TAATCTTGGTGATCTATTGAAGGATCT 943  
DB 1136 TAATCTTGGTGATCTATTGAAGGATCT 1164

RESULT 8  
AAT99556  
JD AAT99556 standard: DNA: 882 BP.  
XX  
XX AAT99556;  
XX  
XX 08-JUN-1998 (first entry)  
XX  
XX Phytolacca insularis antiviral protein gPIP2 gene.  
XX  
XX Antiviral protein: gPIP2 gene; virucide: transgenic plant;  
XX virus resistance; immunocombinator; AIDS; cancer; therapy: ss.  
XX  
XX Phytolacca insularis Nakal.  
XX  
XX Key Location/Qualifiers  
XX CDS 1..882  
XX /tag= a  
XX /transl\_except= (pos:715..717, aa:lle)  
XX  
XX EP808902-A2.  
XX  
XX 26-NOV-1997.  
XX  
XX 30-SEP-1996; 96EP-0307159.  
XX  
XX 22-MAY-1996; 96KR-0017404.  
XX  
XX (JINR-) JIN RO 13D.  
XX  
XX Choi J, Choi K, Choi Y, Hong E, Jin J, Kim C;  
XX Kim M, Lee J, Lee J, Moon Y, Song S, Yun Y;  
XX WPI: 1998-001788/01.  
XX



06/02/2003

Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 915 TAATCTGGTGATCTATTTCAGGATCT 943  
 |||||  
 Db 944 TAATCTGGTGATCTATTTCAGGATCT 972

## RESULT 4

AAZ59220  
 ID AAZ59220 standard; CDNA; 1379 BP.

XX AC AAZ59220;

XX DT 20-APR-2000 (first entry)

XX DE Pokeweed antiviral protein coding sequence spring leaf form.

XX DE Pokeweed antiviral protein spring leaf form; PAP; transgenic plant;  
 resistance; potato virus X; potato virus Y; potato leaf roll virus;  
 tuber; ss.

XX OS Phytolacca americana.

XX PN US6015940-A.

XX PD 18-JAN-2000.

XX PF 07-APR-1992; 92US-0865169.

XX PR 07-APR-1992; 92US-0865169.

XX PA (MONS ) MONSANTO CO.

XX PI Kaniewski WK, Turner NE, Lodge JK;

XX DR WPI: 2000-126326/11.

XX PT Production of transgenic potato plants or tubers expressing pokeweed  
 antiviral protein which are resistant to potato virus X or Y.

XX PS Claim 6; Fig 4; 30pp; English.

XX CC This is the coding sequence for the spring leaf form of the pokeweed  
 antiviral protein (PAP) which is used to generate transgenic potato  
 plants. PAP is able to confer resistance to infection by potato virus  
 X (PVX) or potato virus Y (PVY) or potato leaf roll virus (PLRV) in  
 the potato plant or tuber expressing PAP.

XX SO Sequence 1379 BP; 488 A; 233 C; 269 G; 389 T; 0 other;

Query Match 3.1%; Score 29; DB 21; Length 1379;

Best Local Similarity 100.0%; Pred. No. 0.00022;

Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 915 TAATCTGGTGATCTATTTCAGGATCT 943  
 |||||  
 Db 1136 TAATCTGGTGATCTATTTCAGGATCT 1164

## RESULT 5

AAZ59221  
 ID AAZ59221 standard; CDNA; 1379 BP.

XX AC AAZ59221;

XX DT 20-APR-2000 (first entry)

XX DE Variant pokeweed antiviral protein spring leaf form coding sequence.

XX DE Pokeweed antiviral protein spring leaf form; PAP; transgenic plant;  
 resistance; potato virus X; potato virus Y; potato leaf roll virus;  
 tuber; ss.

XX

OS Phytolacca americana.

XX PN US6015940-A.

XX PD 18-JAN-2000.

XX PF 07-APR-1992; 92US-0865169.

XX PR 07-APR-1992; 92US-0865169.

XX PA (MONS ) MONSANTO CO.

XX PI Kaniewski WK, Turner NE, Lodge JK;

XX DR WPI: 2000-126326/11.

XX PT Production of transgenic potato plants or tubers expressing pokeweed  
 antiviral protein which are resistant to potato virus X or Y.

XX PS Claim 7; Fig 5; 30pp; English.

XX CC This is the coding sequence for a variant spring leaf form of the  
 pokeweed antiviral protein (PAP) which is used to generate transgenic  
 potato plants. PAP is able to confer resistance to infection by potato  
 virus X (PVX) or potato virus Y (PVY) or potato leaf roll virus (PLRV)  
 in the potato plant or tuber expressing PAP. PAP varies from PAP  
 (AAZ59220) by mutations L20R and Y49H.

XX SO Sequence 1379 BP; 489 A; 234 C; 269 G; 387 T; 0 other;

Query Match 3.1%; Score 29; DB 21; Length 1379;

Best Local Similarity 100.0%; Pred. No. 0.00022;

Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 915 TAATCTGGTGATCTATTTCAGGATCT 943

|||||

Db 1136 TAATCTGGTGATCTATTTCAGGATCT 1164

## RESULT 6

AAZ45197  
 ID AAZ45197 standard; DNA; 1379 BP.

XX AC AAZ45197;

XX DT 29-FEB-2000 (first entry)

XX DE Wild-type pokeweed antiviral protein (PAP) nucleotide sequence.

XX DE Pokeweed antiviral protein; PAP; PAP II; antifungal; transgenic plant;  
 ribosome inhibiting protein; RIB; cereal crop; viral resistance; PVX;  
 potato virus X; cucumber mosaic virus; CMV; SS;  
 tomato yellow leaf curl virus.

XX OS Phytolacca americana.

XX FH Key Location/Qualifiers

XX FT CDS 225..1166

XX FT /\*tag=

XX FT /product= PAP

XX FT /note= "Pokeweed antiviral protein"

XX PN WO9960843-A1.

XX PD 02-DEC-1999.

XX PF 21-MAY-1999; 99WO-US11301.

XX PR 22-MAY-1998; 98US-0086374.

XX PA (RUTF ) UNIV ROTTERS STATE NEW JERSEY.

XX PI Turner NE, Wang P;

01-JUN-1993.  
XX  
XX  
PF 20-NOV-1991: 91JP-0329672.  
XX  
PR 20-NOV-1991: 91JP-0329672.  
XX  
XX (NTSB ) JAPAN TOBACCO INC.  
XX  
XX WPI: 1993-211306/26.  
DR P-PSDB: AAR37345.  
XX  
XX New pokeweed antiviral protein (PAP) with similar activity to  
PT ricin - used to treat cancer and as an agricultural chemical  
XX  
XX Claim 2: Page 11-13: 14pp: Japanese.  
XX  
XX PAP has a similar activity to ricin, i.e. inhibits protein synthesis.  
CC The protein may be obtained all year round by recombinant DNA  
CC techniques. PAP can be used partic. against cancer and as an  
CC agricultural chemical.  
CC Total mRNA, is extracted from the seeds, leaves and roots of  
CC pokeweed and used to prepare cDNA using PCR. The resultant cDNA is  
CC used to prepare two DNA fractions, which are introduced into a  
CC cloning vector EMBL3 and then into host E.coli PLK-17 (P2) to  
CC produce PAP.  
XX  
XX Sequence 2472 BP; 864 A; 375 C; 463 G; 770 T; 0 other:  
SQ  
Query Match 3.4%; Score 32; DB 14; Length 2472;  
Best Local Similarity 100.0%; Pred. No. 6.8e-06;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Caps 0;  
Qy 577 GTAGCCATCCAAATGGTTTCAGAGCGCG 608  
Db 1587 GTAGCCATCCAAATGGTTTCAGAGCGCG 1618  
SRNT  
RESULT 2  
AAQ56672  
ID AAQ56672 standard; cDNA; 1195 BP.  
XX  
XX AC AAQ56672:  
XX  
XX 04-SEP-1994 (first entry)  
XX  
XX Sequence of Phytolacca antiviral protein (PAP) cDNA.  
DE  
XX Antiviral protein; PAP; virus-resistance; transgenic plant; ss.  
XX  
XX Phytolacca americana L.  
XX  
XX Key Location/Qualifiers  
FT CDS 33..974  
FT /\*tag- a  
XX  
XX EP585554-A.  
PN  
XX  
XX 09-MAR-1994.  
PD  
XX  
XX 30-JUN-1993: 93EP-0110445.  
PF  
XX  
XX 16-AUG-1992: 92KR-0014895.  
PR  
XX  
XX (JINR-) JIN RO LTD.  
PA  
XX (SHIN-) SHINRO KK.  
PA  
XX Choi K, Jeong H, Jeong H, Kim M, Lee K, Moon Y;  
PI Na B;  
PT  
XX  
XX WPI: 1994-076002/10.  
DR P-PSDB: AAR48548.  
XX  
XX Expression vector for phytolacca antiviral protein - used for  
PT

PT producing transgenic virus-resistant plants and for producing the  
PT antiviral agent  
XX  
XX Disclosure: Fig 1: 15pp: English.  
XX  
XX To isolate PAP gene, total cellular mRNA was purified from leaves of  
CC Phytolacca americana L. Obtd. in Korea. A cDNA library was  
CC constructed. The PAP gene was selected by immunoscreening employing  
CC anti-PAP antibody. A deletion mutant was prepd. from the isolated  
CC PAP gene, and the DNA sequence of the PAP genome was determined.  
XX  
XX Sequence 1195 BP; 420 A; 201 C; 230 G; 344 T; 0 other:  
SQ  
Query Match 3.1%; Score 29; DB 15; Length 1195;  
Best Local Similarity 100.0%; Pred. No. 0.00022;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Caps 0;  
Qy 915 TAATCTTGGTGCATCTATTTCAGAGGATCT 943  
Db 944 TAATCTTGGTGCATCTATTTCAGAGGATCT 972  
RESULT 3  
AAQ81457  
ID AAQ81457 standard; cDNA; 1195 BP.  
XX  
XX AC AAQ81457:  
XX  
XX 25-AUG-1995 (first entry)  
XX  
XX Phytolacca antiviral protein (PAP) cDNA.  
DE  
XX Antiviral protein; vector pMJ12; KCCM 10037; PAP; ss.  
XX  
XX Phytolacca americana L.  
XX  
XX A09350642-A.  
PN  
XX  
XX 19-JAN-1995.  
PD  
XX  
XX 11-NOV-1993: 93AU-0050642.  
PF  
XX  
XX 02-JUL-1993: 93KR-0012360.  
PR  
XX  
XX (JINR-) JIN RO LTD.  
PA  
XX  
XX Hong-seob J, Kwan-ho L, Kyu-whan C, Man-keun K;  
PI Young-ho M, Choi K, Jeon H, Kim M, Lee K, Moon Y;  
PI  
XX WPI: 1995-067518/10.  
DR  
XX  
XX Recombinant vector for producing Phytolacca anti-viral protein  
PT and transformed E. coli useful for making immunoconjugates for  
PT treatment of AIDS  
XX  
XX Claim 1: Fig 1: 27pp: English.  
XX  
XX Total cellular mRNA from leaves of P. americana was used to produce  
CC a cDNA library and this screened with anti-PAP antibody raised in  
CC rabbits against purified PAP. Inserts were isolated from 2 clones  
CC and sequenced to identify a 1195 ORF (AAQ81457) that encodes a 313 AA  
CC PAP including a 22 AA signal peptide. The mature PAP gene was  
CC subjected to PCR amplification using primers AAQ81458 and AAQ81459. The  
CC amplification product was cut with Hind III and inserted into the  
CC commercial FLAG (RTM) vector cut with the same enzyme to form pMJ12.  
CC pMJ12 is deposited with the Korean Collection of Culture and  
CC Microorganism (KCCM), an International Depository Authority, on  
CC June 30 1993, ad deposition No. KCCM 10037, and claimed.  
XX  
XX Sequence 1195 BP; 420 A; 203 C; 229 G; 343 T; 0 other:  
SQ  
Query Match 3.1%; Score 29; DB 16; Length 1195;  
Best Local Similarity 100.0%; Pred. No. 0.00022;

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 2, 2003, 01:16:18 : Search time 292 seconds  
(without alignments)  
7288.145 Million cell updates/sec

Title: US-09-978-274A-1  
Perfect score: 945  
Sequence: 1 atgaaggtgatgtttagt.....atctatttgaaggtattctaa 945

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 1125999159 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N\_Geneseq\_101002.\*

- 1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*
- 2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*
- 3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*
- 4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*
- 5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*
- 6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*
- 7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*
- 8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*
- 9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*
- 10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*
- 11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*
- 12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*
- 13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*
- 14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*
- 15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*
- 16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*
- 17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*
- 18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*
- 19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*
- 20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*
- 21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*
- 22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*
- 23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*
- 24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	32	3.4	2472	14 AAQ43967	Pokeweed antiviral
2	29	3.1	1195	15 AAQ56672	Sequence of Phytol
3	29	3.1	1195	16 AAQ81457	Phytolacca antivir
4	29	3.1	1379	21 AA259220	Pokeweed antiviral
5	29	3.1	1379	21 AA259221	Pokeweed antiviral
6	29	3.1	1379	21 AA259221	Variant pokeweed a
7	29	3.1	1379	21 AA245197	Wild-type pokeweed
8	23	2.4	882	19 AAC87929	P. americana pokew
9	22	2.3	918	15 AAQ64893	Phytolacca insular

10	22	2.3	918	16 AAT04782	DNA pJM201 fragme
11	22	2.3	2369	19 AAT99557	Phytolacca insular
12	22	2.3	2369	20 ABA96543	Phytolacca insular
13	22	2.1	1956	20 AAX81493	DNA sequence of cr
14	22	2.1	1956	20 AAX81495	DNA sequence of cr
15	20	2.1	1956	20 AAX81496	DNA sequence of cr
16	20	2.1	1959	20 AAX81532	DNA sequence of cr
17	20	2.1	1959	20 AAX81528	DNA sequence of cr
18	20	2.1	1959	20 AAX81528	DNA sequence of cr
19	20	2.1	1959	20 AAX81498	DNA sequence of cr
20	20	2.1	1959	20 AAX81499	DNA sequence of cr
21	20	2.1	1959	20 AAX81500	DNA sequence of cr
22	20	2.1	1959	20 AAX81501	DNA sequence of cr
23	20	2.1	1959	20 AAX81481	DNA sequence of cr
24	20	2.1	1959	20 AAX81482	DNA sequence of cr
25	20	2.1	1959	20 AAX81483	DNA sequence of cr
26	20	2.1	1959	20 AAX81484	DNA sequence of cr
27	20	2.1	1959	20 AAX81485	DNA sequence of cr
28	20	2.1	1959	20 AAX81486	DNA sequence of cr
29	20	2.1	1959	20 AAX81487	DNA sequence of cr
30	20	2.1	1959	20 AAX81488	DNA sequence of cr
31	20	2.1	1959	20 AAX81489	DNA sequence of cr
32	20	2.1	1959	20 AAX81490	DNA sequence of cr
33	20	2.1	1959	20 AAX81491	DNA sequence of cr
34	20	2.1	1959	20 AAX81492	DNA sequence of cr
35	20	2.1	1959	20 AAX81494	DNA sequence of cr
36	20	2.1	1959	20 AAX81470	DNA sequence of cr
37	20	2.1	1959	20 AAX81471	DNA sequence of cr
38	20	2.1	1959	20 AAX81472	DNA sequence of cr
39	20	2.1	1959	20 AAX81473	DNA sequence of cr
40	20	2.1	1959	20 AAX81474	DNA sequence of cr
41	20	2.1	1959	20 AAX81475	DNA sequence of cr
42	20	2.1	1959	20 AAX81476	DNA sequence of cr
43	20	2.1	1959	20 AAX81477	DNA sequence of cr
44	20	2.1	1959	20 AAX81478	DNA sequence of cr
45	20	2.1	1959	20 AAX81479	DNA sequence of cr

ALIGNMENTS

RESULT 1  
AAQ43967  
JD AAQ43967 standard; DNA: 2472 BP.  
XX  
AC AAQ43967;  
XX  
DT 09-NOV-1993 (first entry)  
XX  
DE Pokeweed antiviral protein.  
XX  
KW POKEROOT; ricin; protein synthesis inhibitor; cancer;  
XX polymerase chain reaction; PCR; NS.  
OS Phytolacca americana.  
XX  
FH Key Location/Qualifiers  
FT CAAT\_signal 549...552  
FT CAAT\_signal 627...630  
FT TATA\_signal 845...850  
FT sig\_peptide 1014...1085  
FT mat\_peptide 1086...1868  
FT polyA\_signal 2130...2135  
XX JP05137580-A.  
PN

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

DN nucleic - nucleic search, using sw model

Run on: June 2, 2003, 01:15:17 ; Search time 62 Seconds  
(Without alignments)  
4674.347 Million cell updates/sec

Title: US-09-978-274A-1

Perfect score: 945  
Sequence: 1 atgaaggtagctgtgtagt.....atctatttgaaggattctaa 945

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0A  
Maximum Match 100A  
Listing first 45 summaries

Database: Issued Patents\_NA:  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:  
5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq:  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB ID	Description
1	634	67.1	1195	1	US-08-373-858-1	Sequence 1, Appl
2	634	67.1	1195	1	US-08-342-786B-1	Sequence 1, Appl
3	632.4	66.9	1379	1	US-08-500-611-1	Sequence 1, Appl
4	632.4	66.9	1379	2	US-08-500-694-1	Sequence 1, Appl
5	632.4	66.9	1379	3	US-07-865-169-1	Sequence 1, Appl
6	632.4	66.9	1379	3	US-09-005-273-1	Sequence 1, Appl
7	632.4	66.9	1379	5	PCT-US96-11546-1	Sequence 1, Appl
8	630.8	66.6	1379	3	US-08-501-253A-1	Sequence 2, Appl
9	629.2	66.6	1379	3	US-07-865-169-2	Sequence 2, Appl
10	629.2	66.6	1379	3	US-09-005-273-3	Sequence 3, Appl
11	543.8	57.5	918	1	US-08-138-636-1	Sequence 1, Appl
12	543.8	57.5	918	1	US-08-319-622A-1	Sequence 1, Appl
13	543.8	57.5	918	1	US-08-471-564-1	Sequence 1, Appl
14	88.4	9.4	804	2	US-08-356-161-6	Sequence 6, Appl
15	88.4	9.4	804	2	US-08-356-161-7	Sequence 7, Appl
16	88.4	9.4	804	3	US-08-718-904-22	Sequence 22, Appl
17	88.4	9.4	804	3	US-08-718-904-23	Sequence 23, Appl
18	88.4	9.4	804	5	PCT-US93-05702-6	Sequence 6, Appl
19	88.4	9.4	804	5	PCT-US93-05702-7	Sequence 7, Appl
20	88.4	9.4	804	5	PCT-US95-10973A-6	Sequence 6, Appl
21	88.4	9.4	804	5	PCT-US95-10973A-7	Sequence 7, Appl
22	88.2	9.3	774	5	PCT-US91-05766-1	Sequence 1, Appl
23	82	8.7	804	2	US-08-356-161-4	Sequence 4, Appl
24	82	8.7	804	3	US-08-718-904-20	Sequence 20, Appl
25	82	8.7	804	5	PCT-US93-05702-4	Sequence 4, Appl
26	82	8.7	804	5	PCT-US95-10973A-4	Sequence 4, Appl
27	80.4	8.5	804	2	US-08-356-161-3	Sequence 3, Appl

28	80.4	8.5	804	2	US-08-356-161-5	Sequence 5, Appl
29	80.4	8.5	804	3	US-08-718-904-19	Sequence 19, Appl
30	80.4	8.5	804	3	US-08-718-904-21	Sequence 21, Appl
31	80.4	8.5	804	5	PCT-US93-05702-3	Sequence 3, Appl
32	80.4	8.5	804	5	PCT-US93-05702-5	Sequence 5, Appl
33	80.4	8.5	804	5	PCT-US95-10973A-3	Sequence 3, Appl
34	80.4	8.5	804	5	PCT-US95-10973A-5	Sequence 5, Appl
35	79.6	8.4	1233	3	US-08-718-904-80	Sequence 80, Appl
36	73	7.7	771	5	PCT-US95-10973A-59	Sequence 59, Appl
37	73	7.7	813	1	US-07-901-707-11	Sequence 11, Appl
38	73	7.7	813	1	US-07-901-707-57	Sequence 57, Appl
39	73	7.7	813	1	US-07-988-430-11	Sequence 11, Appl
40	73	7.7	813	1	US-07-988-430-57	Sequence 57, Appl
41	73	7.7	813	1	US-08-425-336-11	Sequence 11, Appl
42	73	7.7	813	1	US-08-488-113B-11	Sequence 11, Appl
43	73	7.7	813	1	US-08-477-484B-11	Sequence 11, Appl
44	73	7.7	813	2	US-08-646-360-11	Sequence 11, Appl
45	73	7.7	813	2	US-08-621-803-246	Sequence 246, App

ALIGNMENTS

RESULT 1  
US-08-373-858-1  
Sequence 1, Application US/08373858  
Patent No. 5633155  
GENERAL INFORMATION:  
APPLICANT: Kim, Man-Keun  
APPLICANT: Lee, Kwun-Ho  
APPLICANT: Na, Byoung-Kook  
APPLICANT: Jeong, Han-Seung  
APPLICANT: Choi, Kyu-Whan  
APPLICANT: Moon, Young-Ho  
APPLICANT: Jeon, Hong-Seob  
TITLE OF INVENTION: Expression Vector for Phytolacca  
TITLE OF INVENTION: Antiviral Protein and Process for Preparing Transgenic  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Darby & Darby  
STREET: 805 Third Avenue  
CITY: New York  
STATE: New York  
COUNTRY: US  
ZIP: 10022-7513  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/373.858  
FILING DATE: 18-JAN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Ludwig, Peter  
REGISTRATION NUMBER: 25,351  
REFERENCE/DOCKET NUMBER: 0136/17986-US1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-527-7770  
TELEFAX: 212-753-6237  
TELEX: 236687  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1195 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:

SRNT

Printed 11/17/2004



[illegible]

US-08-500-611-1  
Sequence 1, Application us/08500611  
Patent No. 5756322  
GENERAL INFORMATION:  
APPLICANT: Turner, Nilgün E.  
TITLE OF INVENTION: Pokeweed Antiviral Protein Mutants  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lerner, David, Littenberg, Krumholz & Mentlik  
STREET: 600 South Avenue West  
CITY: Westfield  
STATE: NJ  
COUNTRY: USA  
ZIP: 07090-1497  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/500.611

461	AAATAAAAAACCATCACATTAATGCTTGACGAAACAAATTTGTATGTGATGGTATTTC	521
303	TGATGCTCTTCAAATGGCAATAGTGTCTGTACCATATATTTAATGATATTGCAAGCAATGGA	362
521	TTCTTCTTTTTCGAAACATAAATGCTGTACCATATCTTTAATGATATCTCAGGTACGCA	580
363	ACGCACATCATGTGACAGAAATAGTCTTTGCTCAAGTCTTATAGTTCCTGTGTGCAATTCGTAT	422
581	ACUCCAAGATGCTAGAGACTACTCTTTGCCCAATATGCCAATTTCTCTGTGTAGTAAAGAATC	640
423	TAATCTAGCAATAGCTTATATCGACCATCTGAAAGAAAGACGACAGTAAGTAACATCAAAATCA	482
641	AAACTTTGATAGTCCGATATCCAACTTGGAAATCAAAAGCGGACGTAAATATCAAAATGCA	700
483	AGTCCAAATTTGGCAATTCANATACTCAGCAGATGACATTCGAAAAATCTCTGCAAGTTGATTC	542
701	GCTCCAACTGGCAATTCANATACTCAGCAGATTAATTCGAAAGATTTCTGAAATTCATCTTC	760
543	ATTCCCTCTAAAAACTCAGGCTTTTCTTCTACTCGTAGCCCATCCAAATTTGGTTTCAGACGC	602
761	ATTCACTCAGAAATTCGAAGCGCAATTCCTATTGTGTAGCCATACAAATGGTATTAAGACGC	820











641	AAACCTTTGATGTCGATATCCAAACATTTGGANTCAAAGCGGGAGTAAATTCAGAAGTCA	700
483	ACTCCAAATTTGGGAATTTCAAATACATCGACAGTCACATTTGGAAAAATCTCTGGAGCTTGATTC	542
701	GGTCCAACTGGGAATTTCAAATACTCCACAGTAATATTGGAAAGCATTTCTGGAGCTGATGC	760
543	ATTCCCTGTAAANAACATGAGGCTTTTTTCTACTGTTAGCCATCCAAATGGTTTCAGAGGC	602
761	ATTCACTGGAANAACCCAGGCCGAATTCGTATTGGTAGCCATACAATTGTATCAGAGGC	820
603	AGCGCGATTTCAAGTACATAGAGAACCAGTCAAGACTAATTTTAATGAGAGCTTCTACCC	662
821	AGCAAGATTTCAAGTACATAGAGAACTAGTGAANACTAATTTTACAGAGCATTCACACC	880
663	TGATCCCAAAAGTAATTPAATTTTCGAGGAGAGTGGGCAAAATCTCTCAGGCCAATTCACAA	722
881	TAATCCCAAGTACTTAATTTGCAAGAGACATGGGGTAAGATTTTCAACAGCAATTCATGA	940
723	TGCCAAGAAATGGGGCTTTTACCCAACCACCTTCGAGCTAGTGGATGCCAAGTACCAAGTG	782
941	TGCCAAGAAATGGAGTTTTTACCCAAACCCTTCGAGCTAGTGGATGCCAGTGGTGCCAAAGTG	1000
783	GATAGTCTCTTAGAGTGGATGAATCAATCGTGATGTGGCACCTCCTTAAGTACGTTAATCG	842
1001	GATAGTGTTCAGAGTGGATGAATCAAGCGCTGATGTAGCACCTCTTAACCTACGTTGGTGG	1060
843	AACTGTGTCAGAACCTT---ACCAAAATGCCATGTTTCTCTCAAGTTATATTTCTACTTA	899
1061	GAGCTGTTCAGAACCTTATAACCAAAATGCCATGTTTCTCTCAACTTATATGCTACTTA	1120
900	TTATAATTAATATGCTAATCTTTGGTGATCTATTTGAAGGATTCATA	945
1121	TTATAATTAACATGGTTAATCTTTGGTGATCTATTTGAAGGATTCATA	1166





US-08-319-622A-1

HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE: ORGANISM: Phytolacca insularis Nakai  
TISSUE TYPE: leaf  
IMMEDIATE SOURCE: CLONE: PIP

Query Match 57.5%; Score 543.8; DB 1; Length 918;  
Best Local Similarity 77.4%; Pred. No. 4.7e-144;  
Matches 714; Conservative 0; Mismatches 197; Indels 12; Gaps 4;

QY 8 TGATGCTGTAGTCTGCTGAGCTTAATAGCTGCTCATTCCTGCACCAACTTCAACTT 67  
DB 2 TGAAGTGTATGCTGCTGAGCAATATCATGATGCTCATCTTCGCACCAACTTCACTT 61

QY 68 GTGCCATAATACGATCACCTTTGATGCTGGAATGCCACCACTTAACAAATATGCCACT 127  
DB 62 GGGCGTGAATACCATCATCTACCATGTTGGAAGTACCACCATTAGAACTATGCAACTT 121

QY 128 TTATGGAATCTCTCGTATCAAGCGAAGATCCAAACTAAATGCTATGCATACCAA 187  
DB 122 TTGGATA----CTTCGTACTGAAGCGGAAGATGCC--AAGTTATGCTATGGAATACCAA 175

QY 188 TGCTACTGATCTAATTCGACCCCTAAGTACTTATTTGCTTAAGCTCCANGGTGCAACC 247  
DB 176 TCGTCCCAATATTGATCAAAATCAATATCATATTTGCTGAGCTCCAAAGTTCAAAATG 235

QY 248 TAAACCACTTACACTAATGCTGAGACGAAATCACTTATACGTGATGGCTATTCTGATC 307  
DB 236 AAGAGGCATCACACTAATGCTGAAGACGAACTAATTTATGATGATGGCTATTCTGATC 295

QY 308 CTTCAATGCAATAGCTGCTTACCATATTTAATGATATTAACAGCAGCGAAGCA 367  
DB 296 CTTACAA--CAATAGGTGCTGCTTCCATCTCTTAAGGCTATCTCAGGTACTGACGGC 352

QY 368 CTGATGTGGAATACTCTTTGCTGCTAGTTCTAGTTCTCTGCTGCTGCTGCTGCTAAT 427  
DB 353 AAGATGTAGAGACTACTCTTTGCCAAATGCCGATCTCGTGTGTTGTTAAACATAAAT 412

QY 428 ACAATAGCTTATATCCGACCATCGAAAGAACAGAGTAATCACTCAAGAAATCAAGTCC 487  
DB 413 ATCATAGTCTATATCCAACTTGGAAATCAAGAGCAGGAGTAATTTCAAGAGTCTGAGTCC 472

QY 488 AATTGGCAATTCAAATACTCAGCAGTACATTCGAAATACTCTGAGGTTGATTCTATCC 547  
DB 473 AACTGGGAATTCGAATACTCGACAGTGGCATTTGGAAGGATTTCTGGAGTACGTCATCA 532

QY 548 CTGTAAAACTGAGGCTTTTCTTCTAGTGTAGCCATCCAAATGGTTTCAGAGCGACGCG 607  
DB 533 CTGAGAACGGAAGTGAATCTCTACTGTAGCCATCAAAATGGTATCAGAGCGACAA 592

QY 608 GATTCAAGTACATAGAACCAAGTCAAGTCAATTTAATAGACATTTACCTCTGATC 667  
DB 593 GATTCAAGTACATAGAGGATCAAGTGAAGAACTAATTTAAGACCACTTCAACCTAATC 652

QY 668 CCAAGTAAATTAATTTGGAGGAGAGTGGGCAAACTCTCTGAGCAATTCACAAATGCCA 727  
DB 653 CCAAGTAAATTAATTTGGAGGAGAGTGGGCAAACTCTCTGAGCAATTCACAAATGCCA 712

QY 728 AAGATGGGCTTTTACCCAACTGAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAG 787  
DB 713 GGAATGGAGTTTACCCAACTCTCTACAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAG 772

QY 788 TTCTTAGAGTGGATGAATCAATCTGATGTGGCACTCTTCAAGTACGTTAATGCAACT 847  
DB 773 TCTTCAAGTGGATGAATCAAGCTGATGTCTACTCTTAACTTAACTTATTTGGAGCT 832

QY 848 GTCAGACAACTT--ACCAATGCGATCTCTCAAGTATTAATTTCTACTTTATTATA 904  
DB 833 GCCAGAGAACTTATAACCAAAATGCCATGTTTCTCAACTTATAATGCTACTTTATTATA 892

QY 905 ATTATATGCTAATCTTGTGTAT 927  
DB 893 ATTATATGCTAATCTTGTGTAT 915

RESULT 13  
US-08-471-564-1  
Sequence 1, Application US/08471564  
Patent No. 5723326  
GENERAL INFORMATION:  
APPLICANT: Moon, Young-Ho  
APPLICANT: Jeon, Hong-Seoh  
APPLICANT: Choi, Kyu-Whan  
APPLICANT: Lee, Kwan Ho  
APPLICANT: Kim, Man Keun  
TITLE OF INVENTION: No. 5723326el Genome Coding Phytolacca Antiviral  
TITLE OF INVENTION: Protein and a Recombinant Expression Vector Therefor  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DARBY & DARBY  
STREET: 805 THIRD AVE.  
CITY: NEW YORK  
STATE: NY  
COUNTRY: USA  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IUM PC Compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/471,564  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/210,396  
FILING DATE:  
APPLICATION NUMBER: US 08/138,636  
FILING DATE: 15-OCT-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Ludwig, S. Peter  
REGISTRATION NUMBER: 25,351  
REFERENCE/DOCKET NUMBER: 0136/18818US  
TELEPHONE: (212) 527-7700  
TELEFAX: (212) 753-6237  
TELEX: 236687  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 918 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Phytolacca insularis  
IMMEDIATE SOURCE:  
CLONE: antiviral protein (PIP)  
US-08-471-564-1

Query Match 57.5%; Score 543.8; DB 1; Length 918;  
Best Local Similarity 77.4%; Pred. No. 4.7e-144;  
Matches 714; Conservative 0; Mismatches 197; Indels 12; Gaps 4;

QY 8 TGATGCTGTAGTCTGCTGAGCTTAATAGCTGCTCATTCCTGCACCAACTTCAACTT 67  
DB 2 TGAAGTGTATGCTGCTGAGCAATATCATGATGCTCATCTTCGCACCAACTTCACTT 61

QY 68 GTGCCATAATACGATCACCTTTGATGCTGGAATGCCACCACTTAACAAATATGCCACT 127  
DB 62 GGGCGTGAATACCATCATCTACCATGTTGGAAGTACCACCATTAGAACTATGCAACTT 121

QY 128 TTATGGAATCTCTCGTATCAAGCGAAGATCCAAACTAAATGCTATGCATACCAA 187  
DB 122 TTGGATA----CTTCGTACTGAAGCGGAAGATGCC--AAGTTATGCTATGGAATACCAA 175

QY 188 TGCTACTGATCTAATTCGACCCCTAAGTACTTATTTGCTTAAGCTCCANGGTGCAACC 247  
DB 176 TCGTCCCAATATTGATCAAAATCAATATCATATTTGCTGAGCTCCAAAGTTCAAAATG 235

QY 248 TAAACCACTTACACTAATGCTGAGACGAAATCACTTATACGTGATGGCTATTCTGATC 307  
DB 236 AAGAGGCATCACACTAATGCTGAAGACGAACTAATTTATGATGATGGCTATTCTGATC 295

QY 308 CTTCAATGCAATAGCTGCTTACCATATTTAATGATATTAACAGCAGCGAAGCA 367  
DB 296 CTTACAA--CAATAGGTGCTGCTTCCATCTCTTAAGGCTATCTCAGGTACTGACGGC 352

QY 368 CTGATGTGGAATACTCTTTGCTGCTAGTTCTAGTTCTCTGCTGCTGCTGCTGCTAAT 427  
DB 353 AAGATGTAGAGACTACTCTTTGCCAAATGCCGATCTCGTGTGTTGTTAAACATAAAT 412

QY 428 ACAATAGCTTATATCCGACCATCGAAAGAACAGAGTAATCACTCAAGAAATCAAGTCC 487  
DB 413 ATCATAGTCTATATCCAACTTGGAAATCAAGAGCAGGAGTAATTTCAAGAGTCTGAGTCC 472

QY 488 AATTGGCAATTCAAATACTCAGCAGTACATTCGAAATACTCTGAGGTTGATTCTATCC 547  
DB 473 AACTGGGAATTCGAATACTCGACAGTGGCATTTGGAAGGATTTCTGGAGTACGTCATCA 532

QY 548 CTGTAAAACTGAGGCTTTTCTTCTAGTGTAGCCATCCAAATGGTTTCAGAGCGACGCG 607  
DB 533 CTGAGAACGGAAGTGAATCTCTACTGTAGCCATCAAAATGGTATCAGAGCGACAA 592

QY 608 GATTCAAGTACATAGAACCAAGTCAAGTCAATTTAATAGACATTTACCTCTGATC 667  
DB 593 GATTCAAGTACATAGAGGATCAAGTGAAGAACTAATTTAAGACCACTTCAACCTAATC 652

QY 668 CCAAGTAAATTAATTTGGAGGAGAGTGGGCAAACTCTCTGAGCAATTCACAAATGCCA 727  
DB 653 CCAAGTAAATTAATTTGGAGGAGAGTGGGCAAACTCTCTGAGCAATTCACAAATGCCA 712

QY 728 AAGATGGGCTTTTACCCAACTGAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAG 787  
DB 713 GGAATGGAGTTTACCCAACTCTCTACAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAG 772

QY 788 TTCTTAGAGTGGATGAATCAATCTGATGTGGCACTCTTCAAGTACGTTAATGCAACT 847  
DB 773 TCTTCAAGTGGATGAATCAAGCTGATGTCTACTCTTAACTTAACTTATTTGGAGCT 832

QY 848 GTCAGACAACTT--ACCAATGCGATCTCTCAAGTATTAATTTCTACTTTATTATA 904  
DB 833 GCCAGAGAACTTATAACCAAAATGCCATGTTTCTCAACTTATAATGCTACTTTATTATA 892

62 GGGCGGTGAATACCATCATCTACCATGTTGGAGAGTACACCATTAGAAGCTATCAACTT 121  
128 TTATGGAATCTTTCGTAATCAAGCGAAGATCCAAACTAAATGCTATGGCATACCAA 187  
122 TTGGATA---CTTCGTAAGCGGCAAGATCC--AAGTTATGCTATGGATACCAA 175  
188 TGCTACCTGATACATAATTCGACCCCTAAGTACTTATTGGTTAGCTCCAAAGGTCAACG 247  
176 TGCTGCCCAATATTGGATCAATCCAAATACATATTGGTTAGCTCCAAAGGTCAATG 235  
248 TAAAGCATTACACTAATCTGAGAGCAATACTTATACGTGATGGCTATCTGATC 307  
236 AAGAGGCATCACACTANTGCTAGAGCAACAATTTATATGATGGCTATCTGATC 295  
308 CTTCAATGGCAATAAGTGTGTTTACCATATTTAATGATATTAAGCACCAGCAACGCA 367  
296 CTTACAA---CAATAGTGTGTTTCCATCTCTTAAAGGCTATCTCAGGTACTGACGG 352  
368 CTGATCTGGAGATCTCTTTCCTCAAGTCTAGTTCCTGCTGTTGCAATCTCCATTAAT 427  
353 AGATGTAGAGACTACTCTTTGCCCCAAATCCGATTCCTGTTGTTGTTAAACATAAAT 412  
428 ACAATAGCTTATATCCGACCATCGAAGAAAGACAGAGTAACTCAAGAAATCAAGTCC 487  
413 ATGATAGTCAATATCCAACTTGAATCAAGAGAGGATTAATCAAGAGTCAAGTCC 472  
488 AATTGGGAATCAAACTACTCAGCAGTGCATTTGGAAAAATCTCTGGAGTTGATTCATCC 547  
473 AACTGGGAATTCGAATCTACTCAGCAGTGGCATTTGGAAGGATTTCTGGAGTCAATCA 532  
548 CTGTAATAACTGAGGCTTTTCTTACTGTTAGTCCATCAAGTAACTTAAATGAGTCTAC 667  
533 CTGAGAGACCGAAGCTGAATTCCTACTGTTAGCCATCAAAATGGTATCAGAGGCAAC 592  
608 GATTCAGTACATAGAGAACCAAGTCAAGTCAAGTAACTTAAATGAGTCTTACCTGATC 667  
593 GATTCAGTACATAGAGAACCAAGTCAAGTCAAGTAACTTAAATGAGTCTTACCTGATC 652  
668 CCAAGTAAATTAATTGGAGGAGAGTGGGGGCAAAATCTCTGAGGCAATTCACAATGCCA 727  
653 CCAAGTAAATTAATTGGAGGAGAGTGGGGGCAAAATCTCTGAGGCAATTCATGTTGCCA 712  
728 AGAATGGGCTTTACCCCAACCATCTGAGTGTGAGTGTGATGATGATGATGATGATGAT 787  
713 GGAATGGAGTTTACCCCAACCATCTGAGTGTGAGTGTGATGATGATGATGATGATGAT 772  
788 TCTTAGAGTGGATGAATCAATCTGATGTGAGTGTGAGTGTGATGATGATGATGATGAT 847  
773 TCTTAGAGTGGATGAATCAATCTGATGTGAGTGTGAGTGTGATGATGATGATGATGAT 832  
848 GTCAGACAACTT---ACCAAAATGCAATCTCTCAAGTGTATTAATTTCTACTTATTATA 904  
833 GTCAGACAACTTAAACCAAAATGCAATCTCTCAAGTGTATTAATTTCTACTTATTATA 892  
905 ATTATATGCTAATCTTGTGTAT 927  
893 ATTATATGCTAATCTTGTGTAT 915

Sequence 6, Application US/08356161  
GENERAL INFORMATION:  
APPLICANT: Leppl, Douglas A.  
APPLICANT: Barthelmy, Isabel  
APPLICANT: Baird, J. Andrew  
APPLICANT: Sosnowski, Barbara A.  
TITLE OF INVENTION: RECOMBINANT PRODUCTION OF  
NUMBER OF INVENTION: SAVORIN-CONTAINING PROTEINS  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED AND BERRY LLP

STREET: 701 Fifth Avenue, 6300 Columbia Center  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/356,161  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/356,161  
FILING DATE: 13-APR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: NO. 5916772tenburg, Carol  
REGISTRATION NUMBER: 39,317  
REFERENCE/DOCKET NUMBER: 760100.404US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 804 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: unknown  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..804  
FEATURE:  
NAME/KEY: mIsC\_feature  
LOCATION: 1..804  
OTHER INFORMATION: /note= "Nucleotide sequence  
OTHER INFORMATION: corresponding to the clone M13 mpl8-G7 in Example 1.B.2."  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 46..804  
OTHER INFORMATION: /product= "Saporin"  
US-08-356-161-6

Query Match 9.4% Score 88.4; DB 2: Length 804;  
Best Local Similarity 49.3%; Pred. No. 1.1e-15;  
Matches 376; Conservative 0; Mismatches 371; Indels 15; Gaps 5;  
QY 47 TTGCTGACCAACTTCAACTGTGTCATATATACCATCACCTTTTGTATGCTGCAATGCA 106  
DB 20 TTTCAGCTTGGACACAACTGATCGGTGCACATCAATCACATTAGATCTAGTAATGCA 79  
QY 107 CCATTACAAATATGCCACCTTTTATGGAATCTCTTGTAAATCAAGCGAAGATCCAAAC 166  
DB 80 CCGCGGTCAATCACTCATCTTTTGGATAAATCCGAACACGCTAAGGATCAAGC 139  
QY 167 TAAATGCTATGGCATACCAATGC---TACCTGATCAATATTCGACCCCTAAGTACTTAT 223  
DB 140 TCAATACGGTGTATACCGACATAGCCGTGATAGGCCACCTTCTAAGAAAATTTCTTA 199  
QY 224 TGGTTAAGCTCCAGGTGCAACCTTAAAGCATTATACACTAATGCTGAGATCAATACCT 283  
DB 200 CAATTAATTTCCAAAGTTC---CCGAGCAAGCTTCTACTTGGCTTAAAGCGATACCT 256  
QY 284 TATAGTGATGGCTTATCTGATCCCTTCAATGCAATAGTGTGCTTACATATATTTA 343  
DB 257 TGTATGCTGCGGTATCTTGCATGATACACGATGTTAATCGGCAATATTTCTA 316  
QY 344 ATGATATTAACAGCAGCAACCCACTGATGTGGAGATACATCTTCTCTCAAGTCTTACT 404  
DB 317 GATCAGAAATTAATTTCCGCGAGTTAACTCCGCTTTTCCAGAGCGCCAGCACTGCAATC 376







Mon Jun 2 10:54:58 2003

time : 65 secs

us-09-978-274a-1.rni

Page 14

06/02/2003

137

SRNT

Printed 11/17/2004

09978274

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

ORF protein - protein search, using sw model

Run on: May 28, 2003, 09:59:45 ; Search time 25 Seconds  
(without alignments)  
520.943 Million cell updates/sec

Title: US-09-978-274A-2

Perfect score: 314

Sequence: 1 MKVNLVVVTLIAWLAAPT.....VIISTYNNYMSNLGDLPEGF 314

Scoring table: OLIGO

Gapop 60.0, Gapext 60.0

Searched: 112892 seqs, 41476328 residues

Word size: 0

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: SwissProt\_40.\*

\*Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	10.2	294	1	RIPA-PHYAM
2	30	9.6	313	1	RIP1-PHYAM
3	28	8.9	261	1	RIPS-PHYAM
4	14	4.5	278	1	RIPP-MIRSA
5	11	3.5	282	1	RIP2-BRDI
6	10	3.2	32	1	RIP2-PHYDI
7	10	3.2	316	1	RIPG-GELMU
8	10	3.2	563	1	RIPG-SAMMI
9	9	2.9	286	1	RIP1-MOMCH
10	9	2.9	286	1	RIP2-MOMBA
11	8	2.5	272	1	FAIL-RHINE
12	7	2.2	130	1	PEDA-THEVO
13	7	2.2	132	1	TVAS-MOUSE
14	7	2.2	258	1	YURA-MYXXA
15	7	2.2	275	1	FA9-RABIT
16	7	2.2	282	1	FA9-RAT
17	7	2.2	300	1	ARGB-BRUHE
18	7	2.2	310	1	RIP2-PHYAM
19	7	2.2	319	1	SYGA-COXBU
20	7	2.2	321	1	HIQA-ARATH
21	7	2.2	384	1	CGB3-CAEEL
22	7	2.2	422	1	YUS3-YEAST
23	7	2.2	429	1	YNO5-MYCTU
24	7	2.2	452	1	FA9-CANFA
25	7	2.2	459	1	FA9-MOUSE
26	7	2.2	507	1	ATPA-MYAZE
27	7	2.2	569	1	AMP2-LYCES
28	7	2.2	607	1	SYRC-YEAST
29	7	2.2	627	1	HEN1-OPSTA
30	7	2.2	743	1	TEF3-HUMAN
31	7	2.2	893	1	PER-PERAM
32	7	2.2	954	1	FLEY-CAUCR
33	7	2.2	990	1	T3RE-SALTY

34 7 2.2 1128 1 PHVA-ORYSA  
35 7 2.2 1522 1 DNA2-YEAST  
36 7 2.2 3591 1 PHAB-BORPE  
37 7 2.2 3707 1 PGBM-MOUSE  
38 6 1.9 19 1 OXLA-OPHUA  
39 6 1.9 62 1 CYEL-MAJKA  
40 6 1.9 70 1 YPO4-STWCO  
41 6 1.9 73 1 VP04-RPAPS  
42 6 1.9 76 1 TXP4-APTSC  
43 6 1.9 76 1 TXP4-APTSC  
44 6 1.9 78 1 GLPE-HUMAN  
45 6 1.9 79 1 CMCL-ETMAC

## ALIGNMENTS

### RESULT 1

RIPA-PHYAM STANDARD: PRI; 294 AA.  
ID RIPA-PHYAM

AC G03464;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-OCT-1993 (Rel. 27, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Antiviral protein alpha precursor (PAP-alpha) (Ribosome-inactivating protein) (rRNA N-glycosidase) (EC 3.2.2.22).  
OS Phytolacca americana (Common pokeweed) (Virginian pokeweed).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; Caryophyllales; Caryophyllaceae; Phytolaccaceae; Phytolacca.  
OX NCBI\_TaxID:3527;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Seed, Leaf, and Root;  
RX MEDLINE=93099240; PubMed=1281438;  
RA Kataoka J., Habuka N., Masuta C., Miyano M., Koizumi A.;  
RT "Isolation and analysis of a genomic clone encoding a pokeweed antiviral protein...";  
RL Plant Mol. Biol. 20:879-886(1992).  
RN [2]  
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).  
RX MEDLINE=95010127; PubMed=7925458;  
RA Ago H., Kataoka J., Tsuge H., Hahuba N., Inagaki E., Noma M., Miyano M.;  
RT "X-ray structure of a pokeweed antiviral protein, coded by a new genomic clone, at 0.23-nm resolution. A model structure provides a suitable electrostatic field for substrate binding.";  
RL Eur. J. Biochem. 225:369-374(1994).  
CC -!- FUNCTION: INHIBITS VIRAL INFECTION OF PLANTS, AND PROTEIN SYNTHESIS IN VITRO. HAS ALSO BEEN SHOWN TO INHIBIT THE REPLICATION OF MAMMALIAN VIRUSES. THE PROTEIN MAY PROVIDE A MEANS OF CELLULAR SUICIDE UPON INVASION BY A VIRUS.  
CC -!- CATALYTIC ACTIVITY: Endonuclease of the N-glycosidic bond at one specific adenosine on the 28S rRNA.  
CC -!- SUBUNIT: MONOMER.  
CC -!- SUBCELLULAR LOCATION: Cell wall.  
CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY. TYPE 1 RIP SUBFAMILY.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcements> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

EMBL: D10600; BAB01451.1;  
PIR: S28421; S28421.  
PDB: LAPA; 3I-JAN-94.  
InterPro: IPR001574; RIP.  
Pfam: PF00161; RIP; 1.  
ProSITE: P500275; SHIGA\_RICIN; 1.









specific adenosine on the 28S rRNA.  
-1- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.  
-1- SIMILARITY: IN THE N-TERMINAL SECTION: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY, TYPE 2 RIP SUBFAMILY.  
-1- SIMILARITY: CONTAINS 2 RICIN B-TYPE LECTIN DOMAINS.  
-----  
This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL: U41299; AAB39475.1;  
PIR: S37382; S37382;  
PIR: S37383; S37383;  
InterPro: IPR001574; RIP;  
InterPro: IPR000772; Ricin\_B\_lectin;  
Pfam: PF00161; RIP; 2;  
Pfam: PF00652; Ricin\_B\_lectin; 6;  
PRINTS: PR00396; SHICARICIN;  
SMART: SM00458; RICIN; 2;  
PROSITE: PS00275; SHIGA-RICIN; 1;  
PROSITE: PS50231; RICIN\_B-LECTIN; 2;  
Plant defense: Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;  
Glycoprotein; Lactin; Signal;  
FT SIGNAL 1 25  
FT CHAIN 26 297 NIGRIN B A CHAIN.  
FT CHAIN 298 563 NIGRIN B B CHAIN.  
FT DOMAIN 305 431 RICIN B-TYPE LECTIN 1.  
FT DOMAIN 434 559 RICIN B-TYPE LECTIN 2.  
FT REPEAT 316 356 1-ALPHA.  
FT REPEAT 357 397 1-BETA.  
FT REPEAT 400 432 1-GAMMA.  
FT REPEAT 445 482 2-ALPHA.  
FT REPEAT 486 524 2-BETA.  
FT REPEAT 527 554 2-GAMMA.  
FT ACT\_SITE 188 188 BY SIMILARITY.  
FT DISULFID 274 302 INTERCHAIN (BY SIMILARITY).  
FT DISULFID 319 338 BY SIMILARITY.  
FT DISULFID 360 377 BY SIMILARITY.  
FT DISULFID 448 463 BY SIMILARITY.  
FT DISULFID 489 506 BY SIMILARITY.  
FT CARBOHYD 221 221 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 368 368 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 376 376 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 483 483 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 537 537 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CONFLICT 39 39 K -> V (IN REF. 2).  
SQ SEQUENCE 563 AA; 62300 MW; F250CBE24621BF14 CRC64;  
-----  
Query Match 3.2%; Score 10; DB 1; Length 563;  
Best Local Similarity 100.0%; Pred. No. 0.029;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
-----  
195 IOMVSEAAEF 204  
|||||||  
183 IOMVSEAAEF 192  
-----  
RIP1\_MOMCH STANDARD; PRT: 286 AA.  
PI6094; P24697;  
01-APR-1990 (Rel. 14, Created)  
01-MAR-1992 (Rel. 21, Last sequence update)  
15-JUN-2002 (Rel. 41, Last annotation update)  
Ribosome-inactivating protein momordin I precursor (rRNA N-glycosidase) (EC 3.2.2.22) (Alpha-momorcharin) (Alpha-NMC).  
Momordica charantia (Bitter melon) (Balsam pear).  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids 1: Cucurbitales: Cucurbitaceae: Momordica.  
OX NCBI\_TaxID=3673;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Seed;  
RX MEDLINE=91159486; PubMed=2001404;  
RA Ho W.K.K., Liu S.C., Shaw P.C., Young H.W., Ng T.B., Chan W.Y.;  
RT "Cloning of the cDNA of alpha-momorcharin: a ribosome inactivating protein.";  
RL Biochim. Biophys. Acta 1088:311-314(1991).  
RN [2]  
RP SEQUENCE OF 24-38.  
RC TISSUE=Seed;  
RX MEDLINE=89326691; PubMed=2753596;  
RA Montecchiucci P.C., Lazzarini A.M., Barbieri L., Stirpe F., Soria M.,  
RA Lappi D.;  
RT "N-terminal sequence of some ribosome-inactivating proteins.";  
RL Int. J. Pept. Protein Res. 33:263-267(1989).  
RN [3]  
RP SEQUENCE OF 24-70.  
RC TISSUE=Seed;  
RX MEDLINE=89005108; PubMed=3262509;  
RA Casellas P., Bussosoy D., Falasca A.I., Barbieri L.,  
RA Guillemot J.C., Ferrara P., Bolognesi A., Cennini P., Stirpe F.;  
RT "Trichokirin, a ribosome-inactivating protein from the seeds of Trichosanthes kirilowii Maximowicz. Purification, partial characterization and use for preparation of immunotoxins.";  
RL Eur. J. Biochem. 176:581-588(1988).  
RN [4]  
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
RX MEDLINE=94356447; PubMed=8075985;  
RA Ren J., Wang Y., Dong Y., Stuart D.I.;  
RT "The N-glycosidase mechanism of ribosome-inactivating proteins implied by crystal structures of alpha-momorcharin.";  
RL StructLett. 342:154-158(1994).  
RN [5]  
RP X-RAY CRYSTALLOGRAPHY (2.16 ANGSTROMS).  
RX MEDLINE=94192822; PubMed=8143869;  
RA Hussain J., Tickle I.J., Wood S.P.;  
RT "Crystal structure of momordin, a type I ribosome inactivating protein from the seeds of Momordica charantia.";  
RL FEBS Lett. 342:154-158(1994).  
RN [6]  
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
RX MEDLINE=95344383; PubMed=7619070;  
RA Huang Q., Liu S., Tang Y., Jin S., Wang Y.;  
RT "Studies on crystal structures, active-centre geometry and depurinating mechanism of two ribosome-inactivating proteins.";  
RL Biochem. J. 309:285-298(1995).  
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one specific adenosine on the 28S rRNA.  
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY. TYPE 1 RIP SUBFAMILY.  
-----  
This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL: X57682; CAA40869.1;  
PIR: S14273; RLPUGG;  
PIR: S16490; S16490;  
PDB: 1A8A; 22-JUN-94.  
PDB: 1A8B; 22-JUN-94.  
PDB: 1A8C; 22-JUN-94.  
PDB: 1MOM; 31-MAY-94.  
PDB: 1MRG; 07-FEB-95.  
PDB: 1MRI; 07-FEB-95.  
GlycoSuiteDB; P16094;

InterPro: IPR001574; RIP.  
Pfam: PF00161; RIP; 1.  
PRINTS: PR00396; SHIGARICIN.  
PROSITE: PS00275; SHIGA\_RICIN; 1.  
Plant defense: Protein synthesis inhibitor; Hydrolase; Toxin; Signal.  
Glycoprotein; 3D-structure.  
SIGNAL 1 23  
CHAIN 24 269 RIBOSOME-INACTIVATING PROTEIN MOMORDIN I.  
PROPEP 270 286 MISSING IN MATURE PROTEIN.  
ACT\_SITE 183 183  
CARBOHYD 250 250 N-LINKED (GLCNAC. . .).  
/FTID-CAR\_000082.  
SEQUENCE 286 AA: 31532 MW: E1B013ABEBC216CF CRC64:

Query Match 2.9% Score 9; DB 1; Length 286;  
Best Local Similarity 100.0%; Pred. No. 0.17;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 200 EAAREFYIE 208  
DB 183 EAAREFYIE 191

RESULT 10  
ID RIP2\_MOMBA STANDARD: PRT: 286 AA.  
AC 29339;  
YT 01-DEC-1992 (Rel. 24, Created)  
YT 01-DEC-1992 (Rel. 24, Last sequence update)  
YT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Ribosome-inactivating protein momordin II precursor (rRNA  
N-glycosidase) (EC 3.2.2.22).  
X Momordica balsamina (Bitter melon) (Balsam pear).  
X Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
X Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
X Eurosoidae I; Cucurbitales; Cucurbitaceae; Momordica.  
NCBI\_TaxID=3672;  
[1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Seed.  
RA MEDLINE=93027170; PubMed=1408771;  
RT Ortigosa M., Better M.;  
X "Momordin II, a ribosome inactivating protein from Momordica  
X balsamina, is homologous to other plant proteins."  
X Nucleic Acids Res. 20:4662-4662(1992).  
X -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
X specific adenosine on the 28S rRNA.  
X -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
X TYPE I RIP SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: Z1175; CAA78166.1;  
PIR: S25360; S25360.  
HSP: P09989; LMRJ.  
InterPro: IPR001574; RIP.  
Pfam: PF00161; RIP; 1.  
PRINTS: PR00396; SHIGARICIN.  
PROSITE: PS00275; SHIGA\_RICIN; 1.  
Plant defense: Protein synthesis inhibitor; Hydrolase; Toxin; Signal.  
SIGNAL 1 23  
CHAIN 24 286 RIBOSOME-INACTIVATING PROTEIN MOMORDIN  
II.  
ACT\_SITE 181 181 BY SIMILARITY  
SEQUENCE 286 AA: 32031 MW: 3889FFFAE6B95986 CRC64:  
Query Match 2.9% Score 9; DB 1; Length 286;

Best Local Similarity 100.0%; Pred. No. 0.17;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 200 EAAREFYIE 208  
DB 181 EAAREFYIE 189

RESULT 11  
ID FAF1\_RHIME STANDARD: PRT: 272 AA.  
AC P58380;  
DT 15-JUN-2002 (Rel. 41, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Enoyl-[acyl-carrier-protein] reductase [NADH] 1 (EC 1.3.1.9) (NADH-  
dependent enoyl-ACP reductase 1).  
GN FABI1 OR R00898 OR SMC00005.  
OS Rhizobium meliloti (Sinorhizobium meliloti).  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Rhizobiaceae; Sinorhizobium.  
OX NCBI\_TaxID=382;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=1021;  
RA MEDLINE=21396507; PubMed=11481430;  
RA Capela D., Barloy-Hubier F., Gouzy J., Bothie G., Ampe F., Ratut J.,  
RA Boistard P., Becker A., Boutry M., Cadieu F., Dreano S., Gloux S.,  
RA Cordie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masny D.,  
RA Pohl T., Portetelle D., Puehler A., Purviller B., Ramsperger U.,  
RA Renard C., Thebaud P., Vandenberg M., Weidner S., Galibert F.;  
X "Analysis of the chromosome sequence of the legume symbiont  
X Sinorhizobium meliloti strain 1021."  
X Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).  
X -1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + NAD(+) -> trans-  
X 2,3-dehydroacyl-[acyl-carrier protein] + NADH.  
X -1- PATHWAY: SECOND REDUCTIVE STEP IN FATTY ACID BIOSYNTHESIS.  
X -1- SUBCELLULAR LOCATION: Inner membrane-associated (by similarity).  
X -1- SIMILARITY: SOME SIMILARITY TO THE SHORT-CHAIN  
X DEHYDROGENASES/REDUCTASES (SDR) FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: AL501785; CAC45470.1;  
InterPro: IPR002198; Adh\_short.  
Pfam: PF00106; adh\_short; 1.  
KW Oxidoreductase; NAD; Fatty acid biosynthesis; Inner membrane;  
KW Complete proteome.  
FT NP\_BIND 14 40 NAD (BY SIMILARITY)  
SQ SEQUENCE 272 AA: 29148 MW: 1E0FAL8A22CDBE36 CRC64;

Query Match 2.5% Score 8; DB 1; Length 272;  
Best Local Similarity 100.0%; Pred. No. 1.8;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 228 LEEKWGI 235  
DB 82 LEEKWGI 89

RESULT 12  
ID PFDA\_THEVO STANDARD: PRT: 130 AA.  
AC Q97BC5;  
DT 15-JUN-2002 (Rel. 41, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)



Prefoldin alpha subunit (GimC alpha subunit).  
PFDA OR TV0531 OR TV05052338.  
Thermoplasma volcanium.  
Archaea: Euryarchaeota; Thermoplasma; Thermoplasmatales;  
Thermoplasmataceae; Thermoplasma.  
NCBI\_TaxID=50339;  
(1)  
SEQUENCE FROM N.A.  
RC STRAIN=CSS1 / DSM 4299 / JCM 9571;  
MEDLINE=20570466; PubMed=11121031;  
RA Kawashima T., Amano N., Koike H., Makino S.-I., Higuchi S.,  
Kawashima-Ohya Y., Matanabe K., Yamazaki M., Kanehori K., Kawamoto T.,  
Nunoshima T., Yamamoto Y., Aramaki H., Makino K., Suzuki M.,  
Archaeal adaptation to higher temperatures revealed by genomic  
sequence of Thermoplasma volcanium.  
Proc. Natl. Acad. Sci. U.S.A. 97:14257-14262(2000).  
CC -!- FUNCTION: Molecular chaperone capable of stabilizing a range of  
proteins. Seems to fulfil an ATP-independent, HSP70-like function  
in archaeal de novo protein folding (By similarity).  
CC -!- SUBUNIT: Heterohexamer of two alpha and four beta subunits (By  
similarity).  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -!- SIMILARITY: BELONGS TO THE PREFOLDIN ALPHA SUBUNIT FAMILY.  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation.  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
EMBL: AP000992; BAB59673.1;  
InterPro: IPR004127; DUF232.  
Pfam: PF02996; DUF232; 1.  
Chaperone, Complete proteome.  
SEQUENCE 130 AA; 14617 MW; 6542126CDCAFA1ED CRC64:  
-----  
Query Match  
Best Local Similarity 100.0%; Pred. No. 9.8;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 166 IQLSSD 172  
DB 37 IQLSSD 43  
-----  
RESULT 13  
TVA3\_MOUSE STANDARD: PRT: 132 AA.  
AC P06323;  
DT 01-JAN-1988 (Rel. 06, Created)  
DT 01-JAN-1988 (Rel. 06, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
CC T-cell receptor alpha chain v region CTL-F3 precursor.  
CC Mus musculus (Mouse).  
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
CC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.  
CC NCBI\_TaxID=10090;  
CC (1)  
CC SEQUENCE FROM N.A.  
MEDLINE=87053852; PubMed=3409068;  
Chou H.S., Behlke M.A., Godambe S.A., Russell J.H., Brooks C.G.,  
Loh D.Y.;  
CC "T cell receptor genes in an allelically diverse CTL clone: implications for  
rearrangement and germline diversity of variable gene segments";  
EMBO J. 5:2149-2155(1986).  
CC PIR: A02014; RWSA3.  
CC InterPro: IPR003599; Ig.  
CC Pfam: PF00047; Ig\_1.  
CC SMART: SM00409; Ig\_1.  
CC T-cell: Receptor; Glycoprotein; Signal.

FT SIGNAL 1 22  
FT CHAIN 23 132  
FT T-CELL RECEPTOR ALPHA CHAIN V REGION (CTL-  
F3).  
FT DOMAIN 23 114  
FT V SEGMENT.  
FT J SEGMENT.  
FT CARBOHYD 43 43  
FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT DISULFID 44 111  
FT NON\_TER 132 132  
SQ SEQUENCE 132 AA; 14747 MW; 4E7DC80D7FEDAA67 CRC64:  
-----  
Query Match  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 88 LMLRRN 94  
DB 15 LMLRRN 21  
-----  
RESULT 14  
YURA\_MYXXA STANDARD: PRT: 258 AA.  
AC P25970;  
DT 01-MAY-1992 (Rel. 22, Created)  
DT 01-MAY-1992 (Rel. 22, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical oxidoreductase in URAA 5' region (EC 1.-.-) (Fragment).  
OS Myxococcus xanthus.  
OC Bacteria; Proteobacteria; delta subdivision; Myxobacteria;  
OC Myxococcales; Cystobacterineae; Myxococcaceae; Myxococcus.  
OX NCBI\_TaxID=34;  
OX (1)  
RP SEQUENCE FROM N.A.  
RC STRAIN=DK1622;  
RX MEDLINE=92112814; PubMed=1730672;  
RA Kimsey H.H., Kaiser D.;  
RT "The orotidine-5'-monophosphate decarboxylase gene of Myxococcus  
xanthus: Comparison to the GMP decarboxylase gene family";  
RL J. Biol. Chem. 267:819-824(1992).  
CC -!- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES  
(SDR) FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation.  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
EMBL: M79325; -; NOT\_ANNOTATED\_CDS.  
DR InterPro: IPR002198; ADH\_short.  
DR Pfam: PF00106; adh\_short; 1.  
DR PROSITE: PS00061; ADH\_SHORT; 1.  
KW Hypothetical protein: Oxidoreductase.  
FT NON\_TER 1 1  
FT NP\_BIND 7 30 NAD OR NADP (BY SIMILARITY).  
FT ACT\_SITE 160 160 BY SIMILARITY.  
SQ SEQUENCE 258 AA; 27367 MW; 7D6B1229D3FE8BF CRC64:  
-----  
Query Match  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 42 TFMESLR 48  
DB 169 TFMESLR 175  
-----  
RESULT 15  
FA9\_RABIT STANDARD: PRT: 275 AA.  
AC P16292;  
ID FA9\_RABIT  
AC P16292;

01-AUG-1990 (Rel. 15, Created)  
01-AUG-1990 (Rel. 15, Last sequence update)  
15-JUN-2002 (Rel. 41, Last annotation update)  
Coagulation factor IX (EC 3.4.21.22) (Christmas factor) (Fragment).  
P9.  
Oryctolagus cuniculus (Rabbit).  
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
CC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
CC NCBI\_TaxID=9986;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90152675; PubMed=2303254;  
RA Sarkar G., Koerber D.D., Sommer S.S.;  
RT "Direct sequencing of the activation peptide and the catalytic domain  
of the factor IX gene in six species.";  
KL Genomics 6:131-143(1990).  
CC -!- FUNCTION: FACTOR IX IS A VITAMIN K-DEPENDENT PLASMA PROTEIN THAT  
CC PARTICIPATES IN THE INTRINSIC PATHWAY OF BLOOD COAGULATION BY  
CC CONVERTING FACTOR X TO ITS ACTIVE FORM IN THE PRESENCE OF CA++  
CC IONS, PHOSPHOLIPIDS, AND FACTOR VIIIA.  
CC -!- CATALYTIC ACTIVITY: Hydrolyzes one Arg-I-Ile bond in factor X to  
CC form factor Xa.  
CC -!- SUBUNIT: FACTOR IX IS ACTIVATED BY FACTOR XIA, WHICH EXCISES THE  
CC ACTIVATION PEPTIDE AND THUS PRODUCES A MOLECULE CONSISTING OF 2  
CC CHAINS HELD TOGETHER BY ONE OR MORE DISULFIDE BONDS.  
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: M26234; AAA31251.1;  
CC HSP: P16293; IPFX.  
CC MEROPS: S01.214;  
CC InterPro: IPR001254; Ser\_protease\_Try.  
CC InterPro: IPR00294; Vilk\_dep\_GLA.  
CC Pfam: PF00089; trypsin; 1.  
CC SMART: SM00020; tryp\_SPC; 1.  
CC PROSITE: PS00011; GLU\_CARBOXYLATION; PARTIAL.  
CC PROSITE: PS00240; TRYPSIN\_DOM; 1.  
CC PROSITE: PS00134; TRYPSIN\_HIS; 1.  
CC PROSITE: PS00135; TRYPSIN\_SER; 1.  
CC Blood coagulation; Plasma; Serine protease; Calcium-binding;  
KW Hydrolase; Glycoprotein.  
FT NON\_TER 1  
FT ACT\_SITE 89 89 CHARGE RELAY SYSTEM.  
FT ACT\_SITE 137 137 CHARGE RELAY SYSTEM.  
FT ACT\_SITE 233 233 CHARGE RELAY SYSTEM.  
FT CARBOHYD 25 25 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 35 35 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 96 96 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 128 128 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT NON\_TER 275 275  
SS SEQUENCE 275 AA; 30776 MW; FE364489AC76BE87 CRC64;  
Query Match 2.2%; Score 7; DB 1; Length 275;  
Best Local Similarity 100.0%; Pred. NO. 20;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
34 NATINKY 40  
1111111  
128 NATINKY 134

Search completed: May 28, 2003, 10:10:15  
Time : 26 secs

Printed 11/17/2004

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

On protein - protein search, using sw model

Run on: May 28, 2003, 10:05:30 ; Search time 85 Seconds  
(without alignments)  
761.162 Million cell updates/sec

Title: US-09-978-274A-2  
Perfect score: 314  
Sequence: 1 MKVNLVVVVTLIAWLIAPT.....VIISTYNNWSNLCOLFEGF 314

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 671580 seqs, 206047115 residues

Word size : 0

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

- 1: SP-archaea.\*
- 2: SP-bacteria.\*
- 3: SP-fungi.\*
- 4: SP-human.\*
- 5: SP-invertebrate.\*
- 6: SP-mammal.\*
- 7: SP-mhc.\*
- 8: SP-organella.\*
- 9: SP-phage.\*
- 10: SP-plant.\*
- 11: SP-rodent.\*
- 12: SP-virus.\*
- 13: SP-vertebrate.\*
- 14: SP-unclassified.\*
- 15: SP-virus.\*
- 16: SP-bacteriap.\*
- 17: SP-archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Rank	Score	Query Match	Length	DB ID	Description
1	314	100.0	314	10 P93444	P93444 phytolacca
2	124	39.5	262	10 Q85946	Q85946 phytolacca
3	30	9.6	237	10 Q9AT83	Q9AT83 phytolacca
4	28	8.9	261	10 Q85947	Q85947 phytolacca
5	21	6.7	313	10 Q94168	Q94168 phytolacca
6	21	6.7	339	10 Q8RYA4	Q8RYA4 phytolacca
7	19	6.1	315	10 Q9XFF8	Q9XFF8 phytolacca
8	14	4.5	305	10 Q8W4U4	Q8W4U4 bougainville
9	11	3.5	136	10 Q8RY69	Q8RY69 gynostemma
10	10	3.2	258	10 Q9S9E4	Q9S9E4 gellonium mu
11	10	3.2	563	10 Q04367	Q04367 sambucus ni
12	10	3.2	563	10 Q945S2	Q945S2 sambucus ni
13	10	3.2	565	10 Q04071	Q04071 sambucus ni
14	10	3.2	566	10 Q04072	Q04072 sambucus ni
15	10	3.2	569	10 P93543	P93543 sambucus ni
16	10	3.2	570	10 Q22415	Q22415 sambucus ni

17	10	3.2	570	10 Q41358	Q41358 sambucus ni
18	10	3.2	592	10 Q8W2E7	Q8W2E7 iris hollan
19	10	3.2	603	10 Q9M653	Q9M653 polygonatum
20	10	3.2	604	10 Q9M654	Q9M654 polygonatum
21	9	2.9	264	10 Q9FSH2	Q9FSH2 monardella c
22	9	2.9	270	10 Q9SA05	Q9SA05 amarantus
23	9	2.9	279	10 Q9ZT25	Q9ZT25 amarantus
24	9	2.9	286	10 Q9FUV7	Q9FUV7 monardella c
25	9	2.9	293	10 Q8VYU0	Q8VYU0 fatiophila cu
26	9	2.9	294	10 Q93Y66	Q93Y66 dianthus ch
27	9	2.9	294	10 Q93Y65	Q93Y65 dianthus ch
28	9	2.9	294	10 Q93Y64	Q93Y64 dianthus ch
29	8	2.5	242	12 Q80897	Q80897 human papil
30	8	2.5	279	10 Q9MSK6	Q9MSK6 chenopodium
31	8	2.5	279	10 Q9AU63	Q9AU63 chenopodium
32	8	2.5	302	10 P93077	P93077 cicerdefendru
33	8	2.5	321	5 Q17149	Q17149 caenorhabdi
34	8	2.5	327	5 Q17150	Q17150 caenorhabdi
35	8	2.5	371	16 Q07018	Q07018 bacillus su
36	8	2.5	458	16 Q92730	Q92730 chlamydia p
37	8	2.5	588	5 P92010	P92010 caenorhabdi
38	8	2.5	953	2 Q85299	Q85299 haemophilus
39	8	2.5	1539	10 Q9LZ75	Q9LZ75 arabidopsis
40	7	2.2	57	2 Q9Z3V1	Q9Z3V1 rhodospirado
41	7	2.2	68	17 Q97XV0	Q97XV0 sulfolobus
42	7	2.2	76	17 Q9RHU3	Q9RHU3 halobacteri
43	7	2.2	88	10 Q8S2B1	Q8S2B1 oryza sativ
44	7	2.2	92	8 Q01307	Q01307 petunia hyb
45	7	2.2	123	3 Q07355	Q07355 saccharomyc

#### ALIGNMENTS

#### RESULT 1

P93444  
ID P93444 PRELIMINARY: PRT: 314 AA.  
AC P93444  
DT 01-MAY-1997 (TREMREL. 03, Created)  
DT 01-MAY-1997 (TREMREL. 03, Last sequence update)  
DT 01-MAY-2002 (TREMREL. 20, Last annotation update)  
DE rRNA -glycosidase precursor (EC 3.2.2.22).  
GN PAP-S.  
OS Phytolacca americana (Common pokeweed) (Virginia pokeweed).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Caryophyllales; Caryophyllaceae; Phytolaccaceae; Phytolacca.  
OX NCBI\_TaxID=3527;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=SEED;  
RX MEDLINE=97263479; PubMed=9109394;  
RA Poyet J., Hoevener A.;  
RT "cDNA cloning of the gene encoding the Antiviral Protein from the  
RT seeds of Phytolacca americana and its expression in E.coli.";  
RL FEBS Lett. 408:97-100(1997).  
CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
CC SPECIFIC ADENOSINE ON THE 2BS RNA.  
CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY.  
CC EMBL: X98079; CAA66702.1; -.  
DR ISSP: 003464; JAPA.  
DR InterPro: IPR001574; RIP.  
DR Pfam: PF00161; RIP; 1.  
DR PRINTS: PR003356; SHIGARICIN.  
DR PROSITE: PS00275; SHIGA-RICIN; 1.  
KW Hydrolase; Signal; Toxin.  
FT SIGNAL 1 24 POTENTIAL.  
FT CHAIN 25 314 POTENTIAL.  
SQ SEQUENCE 314 AA: 35323 MW: 48953CE57789FF9E CMC64:

Query Match 100.0%; Score 314; DB 10; Length 314;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 314; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

09978274

1 MKVNLVVVTLIAHLIAAPTSTCAINTITDAGNATINKYATFMSLRNKAOPKLCY 60  
1 MKVNLVVVTLIAHLIAAPTSTCAINTITDAGNATINKYATFMSLRNKAOPKLCY 60  
61 IPMLPOTNSTPKYLLVKGANLKTITLMLRRNNLYVMGSDPFGNKKRYHIFNDITST 120  
61 IPMLPOTNSTPKYLLVKGANLKTITLMLRRNNLYVMGSDPFGNKKRYHIFNDITST 120  
121 ERTDVENTICSSSSSRVAMSINNSLYPTWTKKAEVNSRNOVOLGTOILSSDGIKISGVD 180  
121 ERTDVENTICSSSSSRVAMSINNSLYPTWTKKAEVNSRNOVOLGTOILSSDGIKISGVD 180  
181 SPVKTEAFLLVLAQVSEAAARFKYIENOVKTNFNRAFPDPKVINLEEKWKGISEAIIH 240  
181 SPVKTEAFLLVLAQVSEAAARFKYIENOVKTNFNRAFPDPKVINLEEKWKGISEAIIH 240  
241 NAKNGALPKLELVDAKGTWIVLRVDEINRVALKLYVNGTCOTTYQNAFMSQVLIISTY 300  
241 NAKNGALPKLELVDAKGTWIVLRVDEINRVALKLYVNGTCOTTYQNAFMSQVLIISTY 300  
301 YNYMSNLGDLFEFG 314  
301 YNYMSNLGDLFEFG 314

## RESULT 2

Q8S946 PRELIMINARY; PRT: 262 AA.  
AC Q8S946: 01-JUN-2002 (TremBLrel. 21, Created)  
DT 01-JUN-2002 (TremBLrel. 21, Last sequence update)  
DE 01-JUN-2002 (TremBLrel. 21, Last annotation update)  
GN PAP-S2 (Fragment).  
OS Phytolacca americana (Common pokeweed) (Virginian pokeweed).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Caryophyllidae; Caryophyllales; Phytolaccaceae; Phytolacca.  
OX NCBI\_TaxID=3527;  
RN (1)  
RP SEQUENCE FROM N.A.  
RA Honjo E., Watanabe K.;  
RT "Cloning of genomic DNA encoding two types of pokeweed antiviral  
RT protein in seeds, PAP-S1 and PAP-S2, and functional comparison of  
RT their recombinant proteins with other PAP isoforms."  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AB071855; BAB86350.1;  
FT NON\_TER 1  
FT NON\_TER 262  
SQ SEQUENCE 262 AA: 29486 MW: AF2D010A73C9D18B CRC64;  
Query Match 39.5%; Score 124; DB 10; Length 262;  
Best Local Similarity 100.0%; Pred. No. 3.6e-125;  
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
161 OVQLGQILSSDGIKISGVDSPVKTEAFLLVLAQVSEAAARFKYIENOVKTNFNRAFY 220  
137 OVQLGQILSSDGIKISGVDSPVKTEAFLLVLAQVSEAAARFKYIENOVKTNFNRAFY 156  
221 PDPKVINLEEKWKGISEAIIHNAKNGALPKLELVDAKGTWIVLRVDEINRVALKLYVN 280  
197 PDPKVINLEEKWKGISEAIIHNAKNGALPKLELVDAKGTWIVLRVDEINRVALKLYVN 256  
281 GTCC 284  
257 GTCC 260

## RESULT 3

Q8ATN3 PRELIMINARY; PRT: 237 AA.  
Q8ATN3:  
OSATN3;  
Q8ATN3;

01-JUN-2001 (TremBLrel. 17, Created)  
01-JUN-2001 (TremBLrel. 17, Last sequence update)  
01-JUN-2002 (TremBLrel. 21, Last annotation update)  
DE RNA\_glycosidase (EC 3.2.2.22) (Fragment).  
GN MPAP.  
OS Phytolacca americana (Common pokeweed) (Virginian pokeweed).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Caryophyllidae; Caryophyllales; Phytolaccaceae; Phytolacca.  
OX NCBI\_TaxID=3527;  
RN (1)  
RP SEQUENCE FROM N.A.  
RA Chen D., Wang X., Zhou G.;  
RT "Pokeweed antiviral protein gene, partial cds."  
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
CC SPECIFIC ADENOSINE ON THE 28S rRNA.  
CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY.  
DR EMBL: AF338910; AAK21951.1;  
DR HSSP: P10297; LOCC  
DR InterPro: IPR001574; RIP.  
DR Pfam: PF00161; RIP; 1.  
DR PRINTS: PRO0396; SHIGARICIN.  
DR PROSITE: PS00275; SHIGA\_RICIN; 1.  
KW Hydrolase; Toxin.  
FT NON\_TER 1  
FT NON\_TER 237  
SQ SEQUENCE 237 AA: 26534 MW: FE4ADCEE03464783 CRC64;  
Query Match 9.6%; Score 30; DB 10; Length 237;  
Best Local Similarity 100.0%; Pred. No. 1.1e-23;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 190 FLVLAIQMVSEAAARFKYIENOVKTNFNRAF 219  
DB 166 FLVLAIQMVSEAAARFKYIENOVKTNFNRAF 195  
RESULT 4  
Q8S947 PRELIMINARY; PRT: 261 AA.  
AC Q8S947: 01-JUN-2002 (TremBLrel. 21, Created)  
DT 01-JUN-2002 (TremBLrel. 21, Last sequence update)  
DE PAP-S1 (Fragment).  
GN PAP-S1.  
OS Phytolacca americana (Common pokeweed) (Virginian pokeweed).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Caryophyllidae; Caryophyllales; Phytolaccaceae; Phytolacca.  
OX NCBI\_TaxID=3527;  
RN (1)  
RP SEQUENCE FROM N.A.  
RA Honjo E., Watanabe K.;  
RT "Cloning of genomic DNA encoding two types of pokeweed antiviral  
RT protein in seeds, PAP-S1 and PAP-S2, and functional comparison of  
RT their recombinant proteins with other PAP isoforms."  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AB071854; BAB86349.1;  
FT NON\_TER 1  
FT NON\_TER 261  
SQ SEQUENCE 261 AA: 29199 MW: D88B998BEFE1F989 CRC64;  
Query Match 8.9%; Score 28; DB 10; Length 261;  
Best Local Similarity 100.0%; Pred. No. 1.8e-21;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 190 FLVLAIQMVSEAAARFKYIENOVKTNFNR 217  
DB 165 FLVLAIQMVSEAAARFKYIENOVKTNFNR 192

RESULT 5  
 09978274  
 0941C8 PRELIMINARY: PRT: 313 AA.  
 0941C8: 01-DEC-2001 (TREMBlrel. 19, Created)  
 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 rRNA -glycosidase (EC 3.2.2.22).  
 PAP.  
 CN Phytolacca acinosa.  
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Caryophyllidae; Caryophyllales; Phytolaccaceae; Phytolacca.  
 OX NCBI\_TaxID=107615;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Peng X., Yuan J., Qiang B.;  
 RT Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 CC SPECIFIC ADENOSINE ON THE 28S RRNA.  
 CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY.  
 DR EMBL: AY049785; RAL15442.1;  
 DR InterPro: IPR001574; RIP.  
 DR Pfam: PF00161; RIP: 1.  
 DR Hydrolase; Toxin.  
 KW Hydrolase; Toxin.  
 SQ SEQUENCE 313 AA: 35059 MW; E478ECD571C17885 CRC64;  
 Query Match 6.7%; Score 21; DB 10; Length 313;  
 Best Local Similarity 100.0%; Pred. No. 7.5e-14;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QV 84 KTITLMLRRNNLYVMGYSDFP 104  
 82 KTITLMLRRNNLYVMGYSDFP 102  
 RESULT 6  
 0941C8 PRELIMINARY: PRT: 339 AA.  
 0941C8: 01-JUN-2002 (TREMBlrel. 21, Created)  
 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Ribosome inactivating protein type 1 precursor.  
 OS Phytolacca americana (Common pokeweed) (virginian pokeweed).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Caryophyllidae; Caryophyllales; Phytolaccaceae; Phytolacca.  
 OX NCBI\_TaxID=3527;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Park S.-W., Lawrence C.B., Linden J.C., Vivanco J.M.;  
 RT "Characterization of a novel ethylene-inducible ribosome-inactivating  
 protein exuded from root cultures of Phytolacca americana";  
 Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 EMBL: AY071928; AAL61546.1;  
 KX Signal.  
 KW SIGNAL.  
 SQ SEQUENCE 339 AA: 37978 MW; 7D47BDC2DEDD905F CRC64;  
 Query Match 6.7%; Score 21; DB 10; Length 339;  
 Best Local Similarity 100.0%; Pred. No. 8e-14;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QV 190 FLVATQMWSEAAAFKFIENQ 210  
 215 FLVATQMWSEAAAFKFIENQ 235  
 RESULT 7  
 0941C8 PRELIMINARY: PRT: 315 AA.  
 0941C8: 01-JUN-2002 (TREMBlrel. 21, Created)  
 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Ribosome inactivating protein type 1 precursor.  
 OS Phytolacca americana (Common pokeweed) (virginian pokeweed).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Caryophyllidae; Caryophyllales; Phytolaccaceae; Phytolacca.  
 OX NCBI\_TaxID=3527;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Park S.-W., Lawrence C.B., Linden J.C., Vivanco J.M.;  
 RT "Characterization of a novel ethylene-inducible ribosome-inactivating  
 protein exuded from root cultures of Phytolacca americana";  
 Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 EMBL: AY071928; AAL61546.1;  
 KX Signal.  
 KW SIGNAL.  
 SQ SEQUENCE 315 AA: 35059 MW; E478ECD571C17885 CRC64;  
 Query Match 4.5%; Score 14; DB 10; Length 305;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-06;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QV 195 IOWVSEAAAFKFIENQ 208  
 186 IOWVSEAAAFKFIENQ 199

01-NOV-1999 (TREMBlrel. 12, Created)  
 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 rRNA -glycosidase (EC 3.2.2.22).  
 GN PIP2.  
 DE Phytolacca insularis.  
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Caryophyllidae; Caryophyllales; Phytolaccaceae; Phytolacca.  
 OX NCBI\_TaxID=63744;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Song S.K., Choi Y., Moon Y.H., Kim S.G., Choi Y.D., Lee J.S.;  
 RT "Systemic induction of a Phytolacca insularis antiviral protein gene  
 by mechanical wounding, jasmonic acid, and abscisic acid";  
 Plant Mol. Biol. 43:439-450(2000).  
 CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 CC SPECIFIC ADENOSINE ON THE 28S RRNA.  
 CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY.  
 DR EMBL: AF141331; RAD32679.1;  
 DR HSP: Q03464; IAPA.  
 DR InterPro: IPR001574; RIP.  
 DR Pfam: PF00161; RIP: 1.  
 DR PRINTS: PR00336; SHIGARICIN.  
 DR PROSITE: PS00275; SHIGA\_RICIN; UNKNOWN\_1.  
 KW Hydrolase; Toxin.  
 SQ SEQUENCE 315 AA: 35728 MW; F85DE21154B5FA15 CRC64;  
 Query Match 6.1%; Score 19; DB 10; Length 315;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-11;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QV 203 RFKYIENOVKTNENAFYF 221  
 204 RFKYIENOVKTNENAFYF 222  
 RESULT 8  
 08W4U4 PRELIMINARY: PRT: 305 AA.  
 08W4U4: 01-MAR-2002 (TREMBlrel. 20, Created)  
 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Bougainvillea spectabilis.  
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Caryophyllidae; Caryophyllales; Nyctaginaceae; Bougainvillea.  
 OX NCBI\_TaxID=146096;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA den Hartog M.F., Tubelli C., Boon L., Heekens S., Ortiz Huissse A.P.,  
 RA de Boer M., Stirpe F.;  
 RT "Cloning and expression of cDNA coding for bouganin: A type I  
 ribosome-inactivating protein from Bougainvillea spectabilis Willd.";  
 Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
 EMBL: AF43416; AAL35962.1;  
 DR InterPro: IPR001574; RIP.  
 DR Pfam: PF00161; RIP: 1.  
 DR PRINTS: PR00336; SHIGARICIN.  
 DR PROSITE: PS00275; SHIGA\_RICIN; UNKNOWN\_1.  
 SQ SEQUENCE 305 AA: 34067 MW; 31505CE91562DCDB CRC64;  
 Query Match 4.5%; Score 14; DB 10; Length 305;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-06;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QV 195 IOWVSEAAAFKFIENQ 208  
 186 IOWVSEAAAFKFIENQ 199

RESULT 9

004367 PRELIMINARY; PRT: 136 AA.  
 ID Q04367  
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Type I ribosome-inactivating protein (fragment).  
 OS Gynostemma pentaphyllum.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Cucurbitales; Cucurbitaceae; Gynostemma.  
 OC NCBI\_TaxID=182084;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Lin Y., Wu Z.J., Lin Q.Y., Xie L.H.;  
 RT "Partial cds of type I ribosome-inactivating protein from Gynostemma  
 pentaphyllum (Thunb.) Makino."  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AV075115; AAL82822.1;  
 FT NON\_TER 136  
 FT NON\_TER 136  
 SQ SEQUENCE 136 AA: 15270 MW: 23436120E24F37C2 CRC64;

Query Match 3.5% Score 11; DB 10; Length 136;  
 Best Local Similarity 100.0%; Pred. No. 0.0023;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Caps 0;

Qy 195 IQMVSEAAAEK 205

Db 89 IQMVSEAAAEK 99

RESULT 10

004367 PRELIMINARY; PRT: 258 AA.  
 ID Q04367  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE rRNA -glycosidase (EC 3.2.2.22).  
 OS Gelonium multiflorum (Euphorbiaceae Himalaya).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Malpighiales; Euphorbiaceae; Gelonium.  
 OC NCBI\_TaxID=3979;  
 RN [1]  
 RP SEQUENCE.  
 RA Rosenblum M.G., Kohr W.A., Beattie K.L., Beattie W.G., Marks W.,  
 RA Toman P.D., Cheung L.;  
 RT "Amino acid sequence analysis, gene construction, cloning, and  
 expression of gelonin, a toxin derived from Gelonium multiflorum."  
 RL J. Interferon Cytokine Res. 15:547-555(1995).  
 DE -!- CATALYTIC ACTIVITY: ENDOLYOLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 SPECIFIC ADENOSINE ON THE 28S RNA.  
 DE -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY.  
 DE HSSP: P09989; INRJ.  
 RN InterPro: IPR001574; RIP.  
 RN Pfam: PF00161; RIP; 1.  
 RN PRINTS: PR00396; SHIGARICIN.  
 RN HydroLase: Toxin.  
 SQ SEQUENCE 258 AA: 28826 MW: 130686E673FAD6B06 CRC64;

Query Match 3.2% Score 10; DB 10; Length 258;  
 Best Local Similarity 100.0%; Pred. No. 0.048;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Caps 0;

Qy 195 IQMVSEAAAEK 204

Db 168 IQMVSEAAAEK 177

RESULT 11

004367 PRELIMINARY; PRT: 503 AA.  
 ID Q04367  
 DT 01-JUL-1997 (TREMBlrel. 04, Created)  
 DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE rRNA -glycosidase precursor (EC 3.2.2.22).  
 OS Sambucus nigra (European elder).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; euasterids II; Dipsacales; Adoxaceae; Sambucus.  
 OC NCBI\_TaxID=4202;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=98112023; PubMed=9450339;  
 RA Van Damme E.J., Roy S., Burte A., Kouge P., Van Leuven F.,  
 RA Peumans W.J.;  
 RT "The major elderberry (Sambucus nigra) fruit protein is a lectin  
 derived from a truncated type 2 ribosome-inactivating protein."  
 RL Plant J. 12:1251-1260(1997).  
 CC -!- CATALYTIC ACTIVITY: ENDOLYOLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 SPECIFIC ADENOSINE ON THE 28S RNA.  
 CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY.  
 DR EMBL: U76524; AAC15886.1;  
 DR HSSP: P02879; 2AAL.  
 DR InterPro: IPR000772; RIGin\_B\_lectin.  
 DR Pfam: PF00652; RIGin\_B\_lectin; 6.  
 DR PRINTS: PR00396; SHIGARICIN.  
 DR SMART: SM00458; RIGIN; 2.  
 DR PROSITE: PS00231; RIGIN\_B\_LECTIN; 2.  
 DR PROSITE: PS00275; SHIGA\_RIGIN; 1.  
 KW HydroLase; Signal; Toxin.  
 FT SIGNAL 1 25  
 FT CHAIN 26 297  
 FT CHAIN 298 563  
 SQ SEQUENCE 563 AA: 62336 MW: 3E02B6C08E796205 CRC64;

Query Match 3.2% Score 10; DB 10; Length 563;  
 Best Local Similarity 100.0%; Pred. No. 0.094;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Caps 0;

Qy 195 IQMVSEAAAEK 204

Db 183 IQMVSEAAAEK 192

RESULT 12

004367 PRELIMINARY; PRT: 563 AA.  
 ID Q04367  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE rRNA -glycosidase (EC 3.2.2.22).  
 OS Sambucus nigra (European elder).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; euasterids II; Dipsacales; Adoxaceae; Sambucus.  
 OC NCBI\_TaxID=4202;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Van Damme E.J.M.;  
 RT "Characterization and cloning of lectins and ribosome-inactivating  
 proteins from Sambucus nigra leaves."  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 CC -!- CATALYTIC ACTIVITY: ENDOLYOLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 SPECIFIC ADENOSINE ON THE 28S RNA.

Query Match 3.2% Score 10; DB 10; Length 563;

Best Local Similarity 100.0%; Pred. No. 0.094;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Caps 0;

Qy 195 IQMVSEAAAEK 204

Db 183 IQMVSEAAAEK 192

CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY.  
DB EMBL: AF409135; AAL04123.1; ..  
DR InterPro: IPR000772; Ricin\_B\_lectin.  
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE rRNA -glycosidase precursor (EC 3.2.2.22).  
DE Pfam: PF00161; RIP: 1.  
DR PROSITE: PS00331; RICIN\_B\_LFECTIN: 2.  
DR PROSITE: PS00375; SHIGA\_RICIN; UNKNOWN\_1.  
KW Hydrolase; Toxin.  
SQ SEQUENCE 563 AA: 62242 MW: 0777CBEDCF33BF10 CRC64:

Query Match 3.2%: Score 10; DB 10; Length 563;

Best Local Similarity 100.0%; Pred. No. 0.094; 0; Indels 0; Gaps 0;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 195 IQMVSEARF 204

|||||

DB 183 IQMVSEARF 192

RESULT 13

OQ4071

ID OQ4071 PRELIMINARY: PRT: 565 AA.

AC OQ4071:

DT 01-JUL-1997 (TREMBLrel. 04, Created)

DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE rRNA -glycosidase precursor (EC 3.2.2.22).

OS Sambucus nigra (European elder).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

OC Asteridae; euasterids II; Dipsacales; Adoxaceae; Sambucus.

NCBI\_TaxID=4202;

[1]

SEQUENCE FROM N.A.

MEDLINE=97236787; PubMed=9079659;

Van Damme E.J.M., Barre A., Rouge P., Van Leuven F., Peumans W.J.;

Isolation and molecular cloning of a novel type 2 ribosome-

inactivating protein with an inactive B chain from elderberry

(Sambucus nigra) bark.;

J. Biol. Chem. 272:8353-8360(1997).

-!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE

SPECIFIC ADENOSINE ON THE 28S RRNA.

-!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY.

CC EMBL: U58357; AAC49672.1; ..

CC HSSP: P02879; 2A1.

InterPro: IPR000772; Ricin\_B\_lectin.

Pfam: PF00652; Ricin\_B\_lectin: 6.

Pfam: PF00161; RIP: 1.

PRINTS: PR00396; SHIGARICIN.

SMART: SM00458; RICIN: 2.

PROSITE: PS00231; RICIN\_B\_LFECTIN: 2.

PROSITE: PS00275; SHIGA\_RICIN: 1.

Hydrolase; Signal; Toxin.

SIGNAL 1 25

CHAIN 26 565

POTENTIAL.

RIBOSOME INACTIVATING PROTEIN, A AND B

CHAIN.

CHAIN 26 283

RIBOSOME INACTIVATING PROTEIN, A CHAIN.

CHAIN 305 565

RIBOSOME INACTIVATING PROTEIN, B CHAIN.

SEQUENCE 565 AA: 62348 MW: CD774CCU2593A26 CRC64:

Query Match 3.2%: Score 10; DB 10; Length 565;

Best Local Similarity 100.0%; Pred. No. 0.094; 0; Indels 0; Gaps 0;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 195 IQMVSEARF 204

|||||

DB 188 IQMVSEARF 197

RESULT 14

OQ4072

ID OQ4072 PRELIMINARY: PRT: 566 AA.  
OQ4072:  
DT 01-JUL-1997 (TREMBLrel. 04, Created)  
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE rRNA -glycosidase precursor (EC 3.2.2.22).  
DE Sambucus nigra (European elder).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Asteridae; euasterids II; Dipsacales; Adoxaceae; Sambucus.  
NCBI\_TaxID=4202;

[1]

SEQUENCE FROM N.A.

MEDLINE=97236787; PubMed=9079659;

Van Damme E.J.M., Barre A., Rouge P., Van Leuven F., Peumans W.J.;

Isolation and molecular cloning of a novel type 2 ribosome-

inactivating protein with an inactive B chain from elderberry

(Sambucus nigra) bark.;

J. Biol. Chem. 272:8353-8360(1997).

-!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE

SPECIFIC ADENOSINE ON THE 28S RRNA.

-!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY.

CC EMBL: U58358; AAC49673.1; ..

CC HSSP: P02879; 2A1.

InterPro: IPR000772; Ricin\_B\_lectin.

InterPro: IPR001574; RIP.

Pfam: PF00652; Ricin\_B\_lectin: 6.

Pfam: PF00161; RIP: 1.

PRINTS: PR00396; SHIGARICIN.

SMART: SM00458; RICIN: 2.

PROSITE: PS00231; RICIN\_B\_LFECTIN: 2.

PROSITE: PS00275; SHIGA\_RICIN: 1.

Hydrolase; Signal; Toxin.

SIGNAL 1 22

POTENTIAL.

RIBOSOME INACTIVATING PROTEIN, A AND B

CHAIN.

CHAIN 23 284

RIBOSOME INACTIVATING PROTEIN, A CHAIN.

CHAIN 306 566

RIBOSOME INACTIVATING PROTEIN, B CHAIN.

SEQUENCE 566 AA: 62733 MW: 6C8FD2C2CC6E005D CRC64;

Query Match 3.2%: Score 10; DB 10; Length 566;

Best Local Similarity 100.0%; Pred. No. 0.095;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 195 IQMVSEARF 204

|||||

DB 189 IQMVSEARF 198

RESULT 15

OQ3543

ID OQ3543 PRELIMINARY: PRT: 569 AA.

AC OQ3543:

DT 01-NOV-1998 (TREMBLrel. 08, Created)

DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Ribosome-inactivating protein SNAI' precursor (RNA N-glycosidase)

(EC 3.2.2.22).

OS Sambucus nigra (European elder).

GN RIP OR \_RIPCSNAI'.

OS Sambucus nigra (European elder).

OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

OC Asteridae; euasterids II; Dipsacales; Adoxaceae; Sambucus.

NCBI\_TaxID=4202;

[1]

SEQUENCE FROM N.A., AND SEQUENCE OF 29-38 AND 304-309.

TISSUE-BARK;

MEDLINE=97325989; PubMed=9183001;

van Damme E.J.M., Roy S., Barre A., Gitorres L., Mostafapour K.,

Rouge P., Van Leuven F., Gires T., Goldstein I.J., Peumans W.J.;

"Elderberry (Sambucus nigra) bark contains two structurally different

Neu5Ac(alpha2,6)Gal/GalNAc-binding type 2 ribosome-inactivating

proteins.;"

Eur. J. Biochem. 245:648-655(1997).  
 FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN  
 SYNTHESIS. THIS IS ACHIEVED THROUGH THE CATALYTIC INACTIVATION OF  
 60S RIBOSOMAL SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF  
 28 S RNA (BY SIMILARITY).  
 FUNCTION: THE B CHAIN IS RESPONSIBLE FOR CELL AGGLUTINATION  
 (LECTIN ACTIVITY). AGGLUTINATION IS INHIBITED BY  
 NEUSAC(APH2,6) LACTOSE, AND N-LINKED GLYCOPROTEINS SUCH AS FETUIN  
 AND OROSOMUCOID (BY SIMILARITY).  
 CATALYTIC ACTIVITY: ENDOPOLYMERIZATION OF THE N-GLYCOSIDIC BOND AT ONE  
 SPECIFIC ADENOSINE ON THE 28S RNA.  
 SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.  
 DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN  
 CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).  
 SIMILARITY: TO OTHER BACTERIAL AND PLANT RIBOSOME-INACTIVATING  
 PROTEINS. BELONGS TO TYPE 2 RIP.  
 EMBL: U66191; AAC49754.1; --  
 HSP: P02879; 2AAL.  
 InterPro: IPR000772; Ricin\_B\_lectin.  
 InterPro: IPR001574; RIP.  
 Pfam: PF00652; Ricin\_B\_lectin; 6;  
 Pfam: PF00161; RIP; 1;  
 PRINTS: PR00396; SHIGARICIN.  
 SMART: SM00458; RICIN; 2;  
 PROSITE: PS00231; RICIN\_B-LECTIN; 2;  
 PROSITE: PS00275; SHIGA-RICIN; 1;  
 Signal: Hydrolase; Glycosidase; Toxin; Glycoprotein; Lectin.  
 FT SIGNAL 1 28  
 FT CHAIN 29 291 RIBOSOME-INACTIVATING PROTEIN SNAI', A  
 FT PEPTIDE 292 303 LINKER PEPTIDE.  
 FT CHAIN 304 569 RIBOSOME-INACTIVATING PROTEIN SNAI', B  
 FT REPEAT 325 372 CHAIN.  
 FT REPEAT 366 401 1-ALPHA.  
 FT REPEAT 404 436 1-BETA.  
 FT REPEAT 448 488 1-GAMMA.  
 FT REPEAT 492 530 2-ALPHA.  
 FT REPEAT 533 566 2-BETA.  
 FT ACT\_SITE 201 201 2-GAMMA.  
 FT DISULFID 286 311 BY SIMILARITY.  
 FT DISULFID 328 347 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 369 381 BY SIMILARITY.  
 FT DISULFID 451 466 BY SIMILARITY.  
 FT DISULFID 495 512 BY SIMILARITY.  
 FT CARBOHYD 40 40 N-LINKED (GLCNAC...), (POTENTIAL).  
 SQ SEQUENCE 569 AA; 62598 MW; 8289C889A3E1F9AD CRC64;

Query Match 3.2% Score 10; DB 10; Length 569;  
 Best Local Similarity 100.0%; Pred. No. 0.095;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 195 IQMVSEARF 204  
 DE 196 IQMVSEARF 205

Search completed: May 28, 2003, 10:11:48  
 J6 time : 87 secs



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

protein - protein search, using sw model

run on: May 28, 2003, 10:08:25 ; Search time 29 Seconds  
(without alignments)  
318.579 Million cell updates/sec

file: US-09-978-274A-2  
perfect score: 314  
sequence: 1 MKVNLVVVTLIAWLAAPT.....VIISTYNNYNSLGLDFEGF 314

scoring table: OLIGO  
Gapop 60.0, Gapext 60.0

searched: 262574 seqs, 29422922 residues

word size: 0

total number of hits satisfying chosen parameters: 262574

minimum DB seq length: 0  
maximum DB seq length: 2000000000

post-processing: Listing first 45 summaries

database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A-COMB.pap.\*  
2: /cgn2\_6/ptodata/1/iaa/5B-COMB.pap.\*  
3: /cgn2\_6/ptodata/1/iaa/6A-COMB.pap.\*  
4: /cgn2\_6/ptodata/1/iaa/6B-COMB.pap.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS92-COMB.pap.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pap.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

result No.	Score	Query Match	Length	DB	ID	Description
1	30	9.6	313	1	US-08-373-858-2	Sequence 2, Appli
2	30	9.6	313	1	US-08-500-611-2	Sequence 2, Appli
3	30	9.6	313	2	US-08-500-694-2	Sequence 2, Appli
4	30	9.6	313	4	US-09-005-273-2	Sequence 2, Appli
5	30	9.6	313	4	US-09-005-273-4	Sequence 4, Appli
6	30	9.6	313	4	US-08-501-253A-2	Sequence 2, Appli
7	30	9.6	313	5	PCT-US96-11546-2	Sequence 2, Appli
8	28	8.9	261	1	US-07-901-707-9	Sequence 9, Appli
9	28	8.9	261	1	US-07-988-430-9	Sequence 9, Appli
10	28	8.9	261	1	US-08-425-336-9	Sequence 9, Appli
11	28	8.9	261	1	US-08-378-761A-79	Sequence 79, Appli
12	28	8.9	261	1	US-08-485-286-79	Sequence 79, Appli
13	28	8.9	261	1	US-08-488-1138-9	Sequence 9, Appli
14	28	8.9	261	1	US-08-477-484B-9	Sequence 9, Appli
15	28	8.9	261	2	US-08-646-360-9	Sequence 9, Appli
16	28	8.9	261	4	US-08-839-765-9	Sequence 9, Appli
17	28	8.9	261	4	US-09-136-389-9	Sequence 9, Appli
18	28	8.9	261	4	US-09-610-838-9	Sequence 9, Appli
19	28	8.9	261	5	PCT-US92-09487-9	Sequence 9, Appli
20	26	8.3	48	4	US-08-776-059-20	Sequence 20, Appli
21	26	8.3	48	4	US-08-776-059-21	Sequence 21, Appli
22	26	8.3	48	4	US-08-776-059-22	Sequence 22, Appli
23	21	6.7	21	4	US-08-776-059-27	Sequence 27, Appli
24	19	6.1	305	1	US-08-138-636-2	Sequence 2, Appli
25	19	6.1	305	1	US-08-319-622A-2	Sequence 2, Appli
26	19	6.1	305	1	US-08-471-564-2	Sequence 2, Appli
27	14	4.5	21	4	US-08-776-059-25	Sequence 25, Appli

28	14	4.5	47	4	US-08-776-059-15	Sequence 15, Appli
29	14	4.5	250	1	US-07-854-845B-8	Sequence 8, Appli
30	14	4.5	250	1	US-07-901-707-8	Sequence 8, Appli
31	14	4.5	250	1	US-07-988-430-8	Sequence 8, Appli
32	14	4.5	250	1	US-08-425-336-8	Sequence 8, Appli
33	14	4.5	250	1	US-08-378-761A-76	Sequence 76, Appli
34	14	4.5	250	1	US-08-485-286-76	Sequence 76, Appli
35	14	4.5	250	1	US-08-488-1138-8	Sequence 8, Appli
36	14	4.5	250	1	US-08-477-484B-8	Sequence 8, Appli
37	14	4.5	250	2	US-08-646-360-8	Sequence 8, Appli
38	14	4.5	250	4	US-08-839-765-8	Sequence 8, Appli
39	14	4.5	250	4	US-09-136-389-8	Sequence 8, Appli
40	14	4.5	250	4	US-09-610-838-8	Sequence 8, Appli
41	14	4.5	250	5	PCT-US92-09487-8	Sequence 8, Appli
42	21	3.5	32	1	US-08-324-301-4	Sequence 4, Appli
43	11	3.5	282	1	US-08-324-301-15	Sequence 15, Appli
44	10	3.2	251	1	US-07-901-707-2	Sequence 2, Appli
45	10	3.2	251	1	US-07-988-430-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1.  
US-08-373-858-2  
; Sequence 2, Application US/08373858  
; Patent No. 5633155

GENERAL INFORMATION:

APPLICANT: Kim, Man-Keun  
APPLICANT: Lee, Kwan-Ilo  
APPLICANT: Na, Byoung-Kook  
APPLICANT: Jeoung, Han-Seung  
APPLICANT: Choi, Kyu-Whan  
APPLICANT: Moon, Young-Ho  
APPLICANT: Jeon, Hong-Seob  
TITLE OF INVENTION: Expression Vector for Phytolacca  
TITLE OF INVENTION: Antiviral Protein and Process for Preparing Transgenic  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Darby & Darby  
STREET: 805 Third Avenue  
CITY: New York  
STATE: New York  
COUNTRY: US  
ZIP: 10022-7513

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA: US/08/373,858  
APPLICATION NUMBER: US/08/373,858  
FILING DATE: 18-JAN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Ludwig, Peter  
REGISTRATION NUMBER: 25,351  
REFERENCE/DOCKET NUMBER: 0136/17986-US1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-527-7770  
TELEFAX: 212-753-6237  
TELEX: 236687

INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 313 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
ORGANISM: Phytolacca americana



Best Local Similarity 100.0%; Pred. No. 4.9e-21;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GENERAL INFORMATION:  
APPLICANT: Uckun, Fatih  
APPLICANT: Tumer, Nilgun  
TITLE OF INVENTION: Biotherapeutic Agents Comprising  
TITLE OF INVENTION: Recombinant PAP and PAP Mutants  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESS: Merchant & Gould  
STREET: 50 South 7th Street, 3100 No. 6146628 West Center  
CITY: Minneapolis

```

1 CITY: Minneapolis
2
3 STATE: NM
4
5 COUNTRY: US
6
7 ZIP: 55402
8
9 COMPUTER READABLE FORM:
10 MEDIUM TYPE: Floppy disk
11 COMPUTER: IBM PC compatible
12 OPERATING SYSTEM: PC-DOS/MS-DOS
13 SOFTWARE: PatentIn Release #1.0, Version #1.25
14
15 CURRENT APPLICATION DATA:
16 APPLICATION NUMBER: US/08/501.253A
17 FILING DATE: 11-JUL-1995

```

FILING DATE: 11-JUL-1995  
 CLASSIFICATION: 424  
 ATTORNEY/AGENT INFORMATION:  
   NAME: Kettleberger, Denise M.  
   REGISTRATION NUMBER: 33.1924  
   REFERENCE/DOCKET NUMBER: 600.323US01  
 TELECOMMUNICATION INFORMATION:  
   TELEPHONE: 612-312-5300  
   TELEFAX: 612-332-9081  
 INFORMATION FOR SEQ ID NO: 2:  
   SEQUENCE CHARACTERISTICS:  
     LENGTH: 313 amino acids  
     TYPE: amino acid  
     TOPOLOGY: linear  
     MOLECULE TYPE: protein  
 US-08-501-253A-2

```

Query Match          9.6%:  Score 30:  DB 4:  Length 313:
Best local similarity 100.0%:  Pred. No. 4.9e-21:
Matches 30:  Conservative 0:  Mismatches 0:  Indels 0:

QY  190  FLUVAIOMYSEAAAFKYIENOVKTENRAF 219
      ||||| ||||| ||||| ||||| |||||
OB  180  FLUVAIOMYSEAAAFKYIENOVKTENRAF 217

RESULT 7
PCT-US96-11546-2
: Sequence 2: Application PC/TUS9611546
: GENERAL INFORMATION:
: APPLICANT: Turner, Nilqun E.
: TITLE OF INVENTION:  HNAS Encoding Pokeweed Antiviral Protein
: TITLE OF INVENTION:  MuLants
: NUMBER OF SEQUENCES:  2
: CORRESPONDENCE ADDRESS:
: ADDRESSES:  Lerner, David, Littenberg, Krumholz & Mentlik
: STREET: 600 South Avenue West
: CITY: Westfield
: STATE: NJ
: COUNTRY: USA
: ZIP: 07030-1497

```

COUNTRY: USA  
ZIP: 07090-1497

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/11546  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/500,694  
FILING DATE: 11-JUL-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Roley, Shawn P.  
REFERENCE/DOCKET NUMBER: OCIRS 3.0-039  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 908-654-5000  
TELEFAX: 908-654-7866  
TELEX: 139-125  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 313 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US96-11546-2

Query Match 9.6% Score 30; DB 5; Length 313;  
Best Local Similarity 100.0%; Pred. No. 4.9e-21;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 190 FLVVAIQMVSEAAAEKYEINQVKTNFNR 219  
D 188 FLVVAIQMVSEAAAEKYEINQVKTNFNR 217  
RESULT 8  
US-07-901-707-9  
Sequence 9, Application US/07501707  
Patent No. 5376546  
GENERAL INFORMATION:  
APPLICANT: Bernhard, Susan L.  
APPLICANT: Better, Marc D.  
APPLICANT: Carroll, Steve F.  
APPLICANT: Lane, Julie A.  
TITLE OF INVENTION: Materials Comprising and Methods of  
NUMBER OF SEQUENCES: 57  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
ADDRESSEE: Bicknell  
STREET: Two First National Plaza, 20 South Clark  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/901,707  
FILING DATE: 19920619  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5416202and, Greta E.  
REGISTRATION NUMBER: 35302  
REFERENCE/DOCKET NUMBER: 27129/30910

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 346-5750  
TELEFAX: (312) 984-5750  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 261 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-901-707-9  
Query Match 8.9% Score 28; DB 1; Length 261;  
Best Local Similarity 100.0%; Pred. No. 3.5e-19;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 190 FLVVAIQMVSEAAAEKYEINQVKTNFNR 217  
D 165 FLVVAIQMVSEAAAEKYEINQVKTNFNR 192  
RESULT 9  
US-07-988-430-9  
Sequence 9, Application US/07988430  
Patent No. 5416202  
GENERAL INFORMATION:  
APPLICANT: Bernhard, Susan L.  
APPLICANT: Better, Marc D.  
APPLICANT: Carroll, Stephen F.  
APPLICANT: Lane, Julie A.  
APPLICANT: Lei, Shau-Ping  
TITLE OF INVENTION: Materials Comprising and Methods of  
NUMBER OF SEQUENCES: 101  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
ADDRESSEE: Bicknell  
STREET: Two First National Plaza, 20 South Clark  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/988,430  
FILING DATE: 19921209  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5416202and, Greta E.  
REGISTRATION NUMBER: 35302  
REFERENCE/DOCKET NUMBER: 31133  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 346-5750  
TELEFAX: (312) 984-9740  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 261 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-988-430-9

Query Match 8.9% Score 28: DB 1: Length 261:

Best Local Similarity 100.0% Pred. No. 3.5e-19:

Mismatches 28: Conservative 0: Indels 0: Gaps 0:

QY 190 FLVATQMVSEARFKYIENQVKTNR 217

DB 165 FLVATQMVSEARFKYIENQVKTNR 192

RESULT 10

US-08-425-336-9

Sequence 9, Application US/08425336

Patent No. 5621083

GENERAL INFORMATION:

APPLICANT: Better, Marc D.

APPLICANT: Carroll, Stephen F.

APPLICANT: Studolka, Gary M.

TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating

TITLE OF INVENTION: Proteins

NUMBER OF SEQUENCES: 140

CORRESPONDENCE ADDRESS:

ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

STREET: 6300 Sears Tower, 233 South Wacker Drive

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60606-6402

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/425,336

FILING DATE: 18-APR-1995

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/064,691

FILING DATE: 12-MAY-1993

APPLICATION NUMBER: US 07/901,707

FILING DATE: 19-JUN-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/787,567

FILING DATE: 04-NOV-1991

ATTORNEY/AGENT INFORMATION:

NAME: Meyers, Thomas C.

REGISTRATION NUMBER: P-36,989

REFERENCE/DOCKET NUMBER: 31394

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/474-6300

TELEFAX: 312/474-0448

TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 261 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-425-336-9

Query Match 8.9% Score 28: DB 1: Length 261:

Best Local Similarity 100.0% Pred. No. 3.5e-19:

Mismatches 28: Conservative 0: Indels 0: Gaps 0:

QY 190 FLVATQMVSEARFKYIENQVKTNR 217

DB 165 FLVATQMVSEARFKYIENQVKTNR 192

RESULT 11

US-08-378-761A-79

Sequence 79, Application US/08378761A

Patent No. 5635384

GENERAL INFORMATION:

APPLICANT: WALSH, TERENCE A

APPLICANT: HEY, TIMOTHY D

APPLICANT: MORGAN, ALICE ER

TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE

TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF

NUMBER OF SEQUENCES: 81

CORRESPONDENCE ADDRESS:

ADDRESSEE: ANDREA T. BORUCKI

STREET: 9330 ZIONSVILLE ROAD

CITY: INDIANAPOLIS

STATE: IN

COUNTRY: US

ZIP: 46268

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/378,761A

FILING DATE: 26-JAN-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: BORUCKI, ANDREA T

REGISTRATION NUMBER: 33651

REFERENCE/DOCKET NUMBER: 38272B

TELECOMMUNICATION INFORMATION:

TELEPHONE: (317) 337-4846

INFORMATION FOR SEQ ID NO: 79:

SEQUENCE CHARACTERISTICS:

LENGTH: 261 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-378-761A-79

Query Match 8.9% Score 28: DB 1: Length 261:

Best Local Similarity 100.0% Pred. No. 3.5e-19:

Mismatches 20: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 190 FLVATQMVSEARFKYIENQVKTNR 217

DB 165 FLVATQMVSEARFKYIENQVKTNR 192

RESULT 12

US-08-485-286-79

Sequence 79, Application US/08485286

Patent No. 5646026

Patent No. 5646026 5646119

GENERAL INFORMATION:

APPLICANT: WALSH, TERENCE A

APPLICANT: HEY, TIMOTHY D

APPLICANT: MORGAN, ALICE ER

TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE

TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF

NUMBER OF SEQUENCES: 81

CORRESPONDENCE ADDRESS:

ADDRESSEE: ANDREA T. BORUCKI

STREET: 9330 ZIONSVILLE ROAD

CITY: INDIANAPOLIS

STATE: IN

COUNTRY: US

ZIP: 46268

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/485,286

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/378761

FILING DATE: 26-JAN-1995

ATTORNEY/AGENT INFORMATION:

NAME: BORUCKI, ANDREA T.

REGISTRATION NUMBER: 33651

REFERENCE/DOCKET NUMBER: 382728

TELEPHONE: (317) 337-4846

INFORMATION FOR SEQ ID NO: 79:

SEQUENCE CHARACTERISTICS:

LENGTH: 261 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-485-286-79

Query Match 8.9%; Score 28; DB 1; Length 261;

Best Local Similarity 100.0%; Pred. No. 3.5e-19;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 190 FLVLAQMVSSEAAKFYIENQVKTNR 217

DB 165 FLVLAQMVSSEAAKFYIENQVKTNR 192

RESULT 13

US-08-488-1138-9

Sequence 9, Application US/084881138

Patent No. 5744580

GENERAL INFORMATION:

APPLICANT: Better, Marc D.

APPLICANT: Carroll, Stephen F.

APPLICANT: Studnika, Gary M.

TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating

TITLE OF INVENTION: Proteins

NUMBER OF SEQUENCES: 169

CORRESPONDENCE ADDRESS:

ADDRESSEE: McAndrews, Held & Malloy, Ltd.

STREET: 500 West Madison Street, 34th floor

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60661

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/488,1138

FILING DATE: 07-JUN-1995

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/425,336

FILING DATE: 18-APR-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/064,691

FILING DATE: 12-MAY-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/988,430

FILING DATE: 09-DEC-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/901,707

FILING DATE: 19-JUN-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/787,567

FILING DATE: 19-JUN-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/787,567

FILING DATE: 04-NOV-1991

ATTORNEY/AGENT INFORMATION:

NAME: McNicholas, Janet M.

REGISTRATION NUMBER: 32,918

REFERENCE/DOCKET NUMBER: 110220507/200-70.P3.C2A

TELEPHONE: 312/707-8800

TELEFAX: 312/707-9155

TELEX: 650 388-1248

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 261 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-488-1138-9

Query Match 8.9%; Score 28; DB 1; Length 261;

Best Local Similarity 100.0%; Pred. No. 3.5e-19;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 190 FLVLAQMVSSEAAKFYIENQVKTNR 217

DB 165 FLVLAQMVSSEAAKFYIENQVKTNR 192

RESULT 14

US-08-477-4848-9

Sequence 9, Application US/084774848

Patent No. 5756699

GENERAL INFORMATION:

APPLICANT: Better, Marc D.

APPLICANT: Carroll, Stephen F.

APPLICANT: Studnika, Gary M.

TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating

TITLE OF INVENTION: Proteins

NUMBER OF SEQUENCES: 169

CORRESPONDENCE ADDRESS:

ADDRESSEE: McAndrews, Held & Malloy, Ltd.

STREET: 500 West Madison Street, 34th floor

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60661

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/477,4848

FILING DATE: 07-JUN-1995

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/425,336

FILING DATE: 18-APR-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/064,691

FILING DATE: 12-MAY-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/988,430

FILING DATE: 09-DEC-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/901,707

FILING DATE: 19-JUN-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/787,567

FILING DATE: 04-NOV-1991

ATTORNEY/AGENT INFORMATION:

NAME: McNicholas, Janet M.

REGISTRATION NUMBER: 32,918

REFERENCE/DOCKET NUMBER: 110220507/200-70.P3.C2A

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/707-8889  
TELEFAX: 312/707-9155  
TELEX: 650 388-1248

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 261 amino acids

TYPE: amino acids

TOPOLOGY: linear

MOLECULE TYPE: protein

IS-08-477-484B-9

Query Match 8.9% Score 28; DB 1: Length 261;

Best Local Similarity 100.0%; Pred. No. 3.5e-19;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 190 FLVAIQMVSEAAARKYIENOVKTNFNR 217

|||||

Db 165 FLVAIQMVSEAAARKYIENOVKTNFNR 192

# RESULT 15

IS-08-646-360-9

Sequence 9, Application US/08646360

Patent No. 5837491

GENERAL INFORMATION:

APPLICANT: Better, Marc D.

APPLICANT: Carroll, Stephen F.

TITLE OF INVENTION: Studnika, Gary M.

TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating

NUMBER OF SEQUENCES: 173

CORRESPONDENCE ADDRESS:

ADDRESSEE: McAndrews, Held & Malloy, Ltd.

STREET: 500 West Madison Street, 34th floor

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60661

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/646,360

FILING DATE: 13-MAY-1996

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US94/05348

FILING DATE: 12-MAY-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/064,691

FILING DATE: 12-MAY-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/988,430

FILING DATE: 09-DEC-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/901,707

FILING DATE: 19-JUN-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/787,567

FILING DATE: 04-NOV-1991

ATTORNEY/AGENT INFORMATION:

NAME: McNicholas, Janet M.

REGISTRATION NUMBER: 32,918

REFERENCE/DOCKET NUMBER: 200-70.P4

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/707-8889

TELEFAX: 312/707-9155

TELEX: 650 388-1248

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 261 amino acids

TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-646-360-9

Query Match 8.9% Score 28; DB 2: Length 261;

Best Local Similarity 100.0%; Pred. No. 3.5e-19;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 190 FLVAIQMVSEAAARKYIENOVKTNFNR 217

|||||

Db 165 FLVAIQMVSEAAARKYIENOVKTNFNR 192

Search completed: May 28, 2003, 10:13:16

Job time : 30 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

protein - protein search, using sw model

Run On: May 28, 2003, 10:11:56 ; Search time 57 Seconds  
(Without alignments)

546,320 Million cell updates/sec

Title: US-09-978-274A-2

Perfect score: 314

Sequence: 1 MKVMLVVVVTLIAWLAAPT.....VIISTYYNMSLGLDFEGF 314

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0 .

Searched: 375593 seqs, 99172665 residues

Word size : 0

Total number of hits satisfying chosen parameters: 375593

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Published Applications\_AA.\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pcp.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pcp.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US05\_NEW\_PUB.pcp.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pcp.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pcp.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pcp.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pcp.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pcp.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pcp.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pcp.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pcp.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pcp.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pcp.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	314	100.0	314	10	US-09-978-274A-2	Sequence 2, Appli
2	262	83.4	263	10	US-09-978-274A-4	Sequence 4, Appli
3	153	48.7	154	10	US-09-978-274A-6	Sequence 6, Appli
4	109	34.7	110	10	US-09-978-274A-8	Sequence 8, Appli
5	14	4.5	250	9	US-09-792-793A-36	Sequence 36, Appl
6	10	3.2	251	10	US-09-765-527-247	Sequence 247, App
7	10	3.2	293	10	US-09-765-527-259	Sequence 259, App
8	10	3.2	309	10	US-09-765-527-253	Sequence 253, App
9	10	3.2	332	10	US-09-765-527-251	Sequence 251, App
10	8	2.5	69	10	US-09-872-523-52	Sequence 62, Appl
11	8	2.5	70	10	US-09-872-523-60	Sequence 60, Appl
12	8	2.5	140	10	US-09-764-864-147	Sequence 147, Ap
13	7	2.2	247	10	US-09-765-272-170	Sequence 170, App
14	7	2.2	489	9	US-09-738-626-3805	Sequence 3805, Ap
15	7	2.2	554	10	US-09-800-396-4	Sequence 2, Appli
16	7	2.2	554	10	US-09-800-396-4	Sequence 4, Appli
17	7	2.2	563	9	US-09-878-672-3	Sequence 3, Appli
18	7	2.2	563	9	US-09-769-787-27	Sequence 27, Appl
19	7	2.2	563	9	US-09-769-787-28	Sequence 28, Appl

Printed 11/17/2004

20	7	2.2	953	10	US-09-888-615-66	Sequence 66, Appl
21	6	1.9	22	10	US-09-864-761-40447	Sequence 40447, A
22	6	1.9	29	10	US-09-864-761-47020	Sequence 47020, A
23	6	1.9	30	10	US-09-864-761-41595	Sequence 41595, A
24	6	1.9	30	10	US-09-864-761-48395	Sequence 48395, A
25	6	1.9	35	10	US-09-870-893-122	Sequence 122, App
26	6	1.9	35	12	US-10-001-879-197	Sequence 177, App
27	6	1.9	39	10	US-09-864-761-35239	Sequence 35239, A
28	6	1.9	41	10	US-09-864-761-41438	Sequence 41438, A
29	6	1.9	46	10	US-09-864-761-42923	Sequence 42923, A
30	6	1.9	52	10	US-09-864-761-33973	Sequence 33973, A
31	6	1.9	53	10	US-09-864-761-44587	Sequence 44587, A
32	6	1.9	57	12	US-10-001-843-154	Sequence 154, App
33	6	1.9	59	9	US-10-092-154-983	Sequence 983, App
34	6	1.9	59	10	US-09-864-761-46641	Sequence 46641, A
35	6	1.9	59	10	US-09-764-847-983	Sequence 983, App
36	6	1.9	66	9	US-10-277-156-2	Sequence 2, Appli
37	6	1.9	66	10	US-09-867-550-806	Sequence 806, App
38	6	1.9	68	9	US-10-277-156-4	Sequence 4, Appli
39	6	1.9	68	10	US-09-864-761-41886	Sequence 41886, A
40	6	1.9	69	10	US-09-864-761-38641	Sequence 38641, A
41	6	1.9	75	10	US-09-864-761-46336	Sequence 46336, A
42	6	1.9	77	9	US-09-796-692-1417	Sequence 1417, Ap
43	6	1.9	77	9	US-09-796-692-1885	Sequence 1885, Ap
44	6	1.9	77	9	US-10-040-862-1417	Sequence 1417, Ap
45	6	1.9	77	9	US-10-040-862-1885	Sequence 1885, Ap

#### ALIGNMENTS

RESULT 1  
US-09-978-274A-2  
; Sequence 2, Application US/0978274A  
; Patent No. US20020116737A1  
; GENERAL INFORMATION:

; APPLICANT: Thomas, Christopher

; APPLICANT: McPherson, Michael

; APPLICANT: Atkinson, Howard

; APPLICANT: Neelam, Anil

; TITLE OF INVENTION: PLANT CELL DEATH SYSTEM

; FILE REFERENCE: 9341-028

; CURRENT APPLICATION NUMBER: US/09/978,274A

; CURRENT FILING DATE: 2001-10-15

; PRIOR APPLICATION NUMBER: 0025225.4

; PRIOR FILING DATE: 2000-10-14

; NUMBER OF SEQ ID NOS: 32

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 2

; LENGTH: 314

; TYPE: PRT

; ORGANISM: Phytolacca americana

US-09-978-274A-2

Query Match 100.0% Score 314; DB 10; Length 314;  
Best Local Similarity 100.0%; Pred. No. 6.4e-283;  
Matches 314; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MKVMLVVVVTLIAWLAAPTSCAINTITFDAGNATINKYATFMESLRNQAQDKPKCYG	60
DB	1	MKVMLVVVVTLIAWLAAPTSCAINTITFDAGNATINKYATFMESLRNQAQDKPKCYG	60
QY	61	IMPLDPTNSTPKYLLVKLGANLKTITLMLRRNNLVNVCYSDPENGKNCRYHIFNDITST	120
DB	61	IMPLDPTNSTPKYLLVKLGANLKTITLMLRRNNLVNVCYSDPENGKNCRYHIFNDITST	120
QY	121	ERTDVENTLCSSSSSRVAMSINYSLSYPTMEKKAFVNSRNOVOIGIQLLSSHICKISWD	180
DB	121	ERTDVENTLCSSSSSRVAMSINYSLSYPTMEKKAFVNSRNOVOIGIQLLSSHICKISWD	180
QY	181	SFPVKTEAFFLLVALQHVSEAAARKYIENOVKTNNKAFYDPDKVINLEKWKGISAIH	240
DB	181	SFPVKTEAFFLLVALQHVSEAAARKYIENOVKTNNKAFYDPDKVINLEKWKGISAIH	240

009978274

06/02/2003



241 NAKNGALPKLELVDAKGTWIVLRVDEINRDVALLKYVNGTCOTTYONAMESOVLIITY 300  
241 NAKNGALPKLELVDAKGTWIVLRVDEINRDVALLKYVNGTCOTTYONAMESOVLIITY 300  
301 YNYSMLGDLFGCF 314  
301 YNYSMLGDLFGCF 314

RESULT 2  
US-09-978-274A-4  
; Sequence 4, Application US/09978274A  
; Patent No. US20020116737A1  
; GENERAL INFORMATION:  
; APPLICANT: Thomas, Christopher  
; APPLICANT: McPherson, Michael  
; APPLICANT: Atkinson, Howard  
; APPLICANT: Neelam, Anil  
; TITLE OF INVENTION: PLANT CELL DEATH SYSTEM  
; FILE REFERENCE: 9341-028  
; CURRENT APPLICATION NUMBER: US/09/978,274A  
; CURRENT FILING DATE: 2001-10-15  
; PRIOR APPLICATION NUMBER: 0025225.4  
; PRIOR FILING DATE: 2000-10-14  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 263  
; TYPE: PRT  
; ORGANISM: Phytolacca americana  
US-09-978-274A-4

Query Match 83.4% Score 262 DB 10 Length 263;  
Best Local Similarity 100.0% Pred. No. 8.1e-235; Mismatches 0; Indels 0; Gaps 0;  
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 25 INTITFDAGNATINKYATFMSLRNQADPKLCKYGIPLPDNTSTPKYLLVKKLOGANLK 84  
DB 2 INTITFDAGNATINKYATFMSLRNQADPKLCKYGIPLPDNTSTPKYLLVKKLOGANLK 61  
QY 85 TITMLRRNNLYVMGYSDFPNCNKRYHIFNDITSTERTDVENTLCSSSSSRVAMSINYN 144  
DB 62 TITMLRRNNLYVMGYSDFPNCNKRYHIFNDITSTERTDVENTLCSSSSSRVAMSINYN 121  
QY 145 SLYPTMEKKAENVSRNQVQLGIQLISSDGIKISGVDSPVKTEAFFLLVAIQMVSEARF 204  
DB 122 SLYPTMEKKAENVSRNQVQLGIQLISSDGIKISGVDSPVKTEAFFLLVAIQMVSEARF 181  
QY 205 KYIENQVKTNFNRAFYDPKVINLEEKWKISETAIIHNAKNGALPKLELVDAKGTWIVL 264  
DB 182 KYIENQVKTNFNRAFYDPKVINLEEKWKISETAIIHNAKNGALPKLELVDAKGTWIVL 241  
QY 265 RVDEINRDVALLKYVNGTCQTT 286  
DB 242 RVDEINRDVALLKYVNGTCQTT 263

RESULT 3  
US-09-978-274A-6  
; Sequence 6, Application US/09978274A  
; Patent No. US20020116737A1  
; GENERAL INFORMATION:  
; APPLICANT: Thomas, Christopher  
; APPLICANT: McPherson, Michael  
; APPLICANT: Atkinson, Howard  
; APPLICANT: Neelam, Anil  
; TITLE OF INVENTION: PLANT CELL DEATH SYSTEM  
; FILE REFERENCE: 9341-028  
; CURRENT APPLICATION NUMBER: US/09/978,274A  
; CURRENT FILING DATE: 2001-10-15  
; PRIOR APPLICATION NUMBER: 0025225.4  
; PRIOR FILING DATE: 2000-10-14

NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 6  
; LENGTH: 154  
; TYPE: PRT  
; ORGANISM: Phytolacca americana  
US-09-978-274A-6

Query Match 48.7% Score 153 DB 10 Length 154;  
Best Local Similarity 100.0% Pred. No. 4.7e-134; Mismatches 0; Indels 0; Gaps 0;  
Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 25 INTITFDAGNATINKYATFMSLRNQADPKLCKYGIPLPDNTSTPKYLLVKKLOGANLK 84  
DB 2 INTITFDAGNATINKYATFMSLRNQADPKLCKYGIPLPDNTSTPKYLLVKKLOGANLK 61  
QY 85 TITMLRRNNLYVMGYSDFPNCNKRYHIFNDITSTERTDVENTLCSSSSSRVAMSINYN 144  
DB 62 TITMLRRNNLYVMGYSDFPNCNKRYHIFNDITSTERTDVENTLCSSSSSRVAMSINYN 121  
QY 145 SLYPTMEKKAENVSRNQVQLGIQLISSDGIKIS 177  
DB 122 SLYPTMEKKAENVSRNQVQLGIQLISSDGIKIS 154

RESULT 4  
US-09-978-274A-8  
; Sequence 8, Application US/09978274A  
; Patent No. US20020116737A1  
; GENERAL INFORMATION:  
; APPLICANT: Thomas, Christopher  
; APPLICANT: McPherson, Michael  
; APPLICANT: Atkinson, Howard  
; APPLICANT: Neelam, Anil  
; TITLE OF INVENTION: PLANT CELL DEATH SYSTEM  
; FILE REFERENCE: 9341-028  
; CURRENT APPLICATION NUMBER: US/09/978,274A  
; CURRENT FILING DATE: 2001-10-15  
; PRIOR APPLICATION NUMBER: 0025225.4  
; PRIOR FILING DATE: 2000-10-14  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 8  
; LENGTH: 110  
; TYPE: PRT  
; ORGANISM: Phytolacca americana  
US-09-978-274A-8

Query Match 34.7% Score 109 DB 10 Length 110;  
Best Local Similarity 100.0% Pred. No. 2e-93; Mismatches 0; Indels 0; Gaps 0;  
Matches 109; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 178 GVDSPVKTEAFFLLVAIQMVSEARFYIENQVKTNFNRAFYDPKVINLEEKWKISET 237  
DB 2 GVDSPVKTEAFFLLVAIQMVSEARFYIENQVKTNFNRAFYDPKVINLEEKWKISET 61  
QY 238 AIHNAKNGALPKLELVDAKGTWIVLRVDEINRDVALLKYVNGTCQTT 286  
DB 62 AIHNAKNGALPKLELVDAKGTWIVLRVDEINRDVALLKYVNGTCQTT 110

RESULT 5  
US-09-792-793A-36  
; Sequence 36, Application US/09792793A  
; Patent No. US20020168370A1  
; GENERAL INFORMATION:  
; APPLICANT: McDonald, John R.  
; APPLICANT: Coggin, Philip  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING SECONDARY TISSUE DAMAGE  
; TITLE OF INVENTION: OTHER INFLAMMATORY CONDITIONS AND DISORDERS  
; FILE REFERENCE: 25020-601b  
; CURRENT APPLICATION NUMBER: US/09/792,793A  
; CURRENT FILING DATE: 2001-02-22

NUMBER OF SEQ ID NOS: 93  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO 36  
 LENGTH: 250  
 TYPE: PRT  
 ORGANISM: Momordica charantia  
 US-09-792-793A-36

Query Match 4.5% Score 14; DB 9; Length 250;  
 Best Local Similarity 100.0%; Pred. No. 4; le-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 194 AIONVSEAAAFK 207  
 DB 162 AIONVSEAAAFK 175

RESULT 6

US-09-765-527-247  
 Sequence 247, Application US/09765527  
 Patent No. US20020006638A1  
 GENERAL INFORMATION:  
 TITLE OF INVENTION: Methods for Recombinant Microbial Production of Fusion Proteins and BPI-Derived Peptides  
 NUMBER OF SEQUENCES: 265  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
 STREET: 6300 Sears Tower, 233 South Wacker Drive  
 CITY: Chicago  
 STATE: Illinois  
 COUNTRY: United States of America  
 ZIP: 60606-6402

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/765,527  
 FILING DATE: 18-Jan-2001  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/621,803  
 FILING DATE: <Unknown>  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Borun, Michael F.  
 REGISTRATION NUMBER: 25,447  
 REFERENCE/DOCKET NUMBER: 27129/33199  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 312/474-6300  
 TELEFAX: 312/474-0448  
 TELEX: 25-3856

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/765,527  
 FILING DATE: 18-Jan-2001  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/621,803  
 FILING DATE: <Unknown>  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Borun, Michael F.  
 REGISTRATION NUMBER: 25,447  
 REFERENCE/DOCKET NUMBER: 27129/33199  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 312/474-6300  
 TELEFAX: 312/474-0448  
 TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 247:

SEQUENCE CHARACTERISTICS:

LENGTH: 251 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 247:

US-09-765-527-247

Query Match 3.2% Score 10; DB 10; Length 251;

Best Local Similarity 100.0%; Pred. No. 0.21;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 195 IONVSEAAAF 204

DB 161 IONVSEAAAF 170

RESULT 7

US-09-765-527-259

Sequence 259, Application US/09765527

Patent No. US20020006638A1  
 GENERAL INFORMATION:  
 APPLICANT: Better, Marc D.  
 TITLE OF INVENTION: Methods for Recombinant Microbial Production of Fusion Proteins and BPI-Derived Peptides  
 NUMBER OF SEQUENCES: 265  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
 STREET: 6300 Sears Tower, 233 South Wacker Drive  
 CITY: Chicago  
 STATE: Illinois  
 COUNTRY: United States of America  
 ZIP: 60606-6402

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/765,527

FILING DATE: 18-Jan-2001

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/621,803

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Borun, Michael F.

REGISTRATION NUMBER: 25,447

REFERENCE/DOCKET NUMBER: 27129/33199

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/474-6300

TELEFAX: 312/474-0448

TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 259:

SEQUENCE CHARACTERISTICS:

LENGTH: 293 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 259:

US-09-765-527-259

Query Match 3.2% Score 10; DB 10; Length 293;

Best Local Similarity 100.0%; Pred. No. 0.24;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 195 IONVSEAAAF 204

DB 183 IONVSEAAAF 192

RESULT 8

US-09-765-527-253

Sequence 253, Application US/09765527

Patent No. US20020006638A1

GENERAL INFORMATION:

APPLICANT: Better, Marc D.

TITLE OF INVENTION: Methods for Recombinant Microbial Production of Fusion Proteins and BPI-Derived Peptides

NUMBER OF SEQUENCES: 265

CORRESPONDENCE ADDRESS:

ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

STREET: 6300 Sears Tower, 233 South Wacker Drive

CITY: Chicago

STATE: Illinois

COUNTRY: United States of America

ZIP: 60606-6402

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/765,527

US-09-765-527-259

Query Match 3.2% Score 10; DB 10; Length 293;

Best Local Similarity 100.0%; Pred. No. 0.24;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 195 IONVSEAAAF 204

DB 183 IONVSEAAAF 192

09978274

FILING DATE: 18-Jan-2001  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/621,803  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Borun, Michael F.  
REGISTRATION NUMBER: 25,447  
REFERENCE/DOCKET NUMBER: 27129/33199  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6300  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 253:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 309 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 253:  
US-09-765-527-253

Query Match 3.2%: Score 10; DB 10; Length 309;  
Best Local Similarity 100.0%; Pred. No. 0.25; 0: Mismatches 0: Indels 0: Gaps 0:  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0:

QY 195 IQMVSEAAKF 204  
|||||  
DB 183 IQMVSEAAKF 192

RESULT 9  
US-09-765-527-251  
Sequence 251, Application US/09765527  
Patent No. US2002000638A1  
GENERAL INFORMATION:  
APPLICANT: Better, Marc D.  
TITLE OF INVENTION: Methods for Recombinant Microbial Production of Fusion Proteins and BPI-Derived Peptides  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/765,527  
FILING DATE: 18-Jan-2001  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/621,803  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Borun, Michael F.  
REGISTRATION NUMBER: 25,447  
REFERENCE/DOCKET NUMBER: 27129/33199  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6300  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 251:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 332 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 251:  
US-09-765-527-251

Query Match 3.2%: Score 10; DB 10; Length 332;  
Best Local Similarity 100.0%; Pred. No. 0.26; 0: Mismatches 0: Indels 0: Gaps 0:  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0:

QY 195 IQMVSEAAKF 204  
|||||  
DB 183 IQMVSEAAKF 192

RESULT 10  
US-09-872-523-62  
Sequence 62, Application US/09872523  
Patent No. US20020137906A1  
GENERAL INFORMATION:  
APPLICANT: Horvitz, H. Robert  
APPLICANT: Davison, Eva M.  
APPLICANT: Lu, Xiaowei  
TITLE OF INVENTION: A Tumor Suppressor Pathway in C. Elegans  
FILE REFERENCE: 01997/536002  
CURRENT APPLICATION NUMBER: US/09/872,523  
CURRENT FILING DATE: 2001-06-01  
PRIOR APPLICATION NUMBER: US 60/208,802  
PRIOR FILING DATE: 2000-06-02  
NUMBER OF SEQ ID NOS: 78  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 62  
LENGTH: 69  
TYPE: PRT  
ORGANISM: Caenorhabditis elegans  
US-09-872-523-62

Query Match 2.5%: Score 8; DB 10; Length 69;  
Best Local Similarity 100.0%; Pred. No. 4.8; 0: Mismatches 0: Indels 0: Gaps 0:  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0:

QY 305 SNLGDLEF 312  
|||||  
DB 43 SNLGDLEF 50

RESULT 11  
US-09-872-523-60  
Sequence 60, Application US/09872523  
Patent No. US20020137906A1  
GENERAL INFORMATION:  
APPLICANT: Horvitz, H. Robert  
APPLICANT: Davison, Eva M.  
APPLICANT: Lu, Xiaowei  
TITLE OF INVENTION: A Tumor Suppressor Pathway in C. Elegans  
FILE REFERENCE: 01997/536002  
CURRENT APPLICATION NUMBER: US/09/872,523  
CURRENT FILING DATE: 2001-06-01  
PRIOR APPLICATION NUMBER: US 60/208,802  
PRIOR FILING DATE: 2000-06-02  
NUMBER OF SEQ ID NOS: 78  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 60  
LENGTH: 70  
TYPE: PRT  
ORGANISM: Caenorhabditis elegans  
US-09-872-523-60

Query Match 2.5%: Score 8; DB 10; Length 70;  
Best Local Similarity 100.0%; Pred. No. 4.9; 0: Mismatches 0: Indels 0: Gaps 0:  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0:

QY 305 SNLGDLEF 312  
|||||  
DB 43 SNLGDLEF 50

RESULT 12

US-09-764-864-1447  
Sequence 1447, Application US/09764864  
Patent No. US20020132753A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PT223  
CURRENT APPLICATION NUMBER: US/09/764.864  
CURRENT FILING DATE: 2001-01-17  
Prior application data removed - consult PALM or file wrapper  
NUMBER OF SEQ ID NOS: 1792  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1447  
LENGTH: 140  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-764-864-1447

Query Match 2.3% Score 8; DB 10; Length 140;  
Best Local Similarity 100.0%; Pred. No. 8.9;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Caps 0;

Y 17 AAPTSTCA 24  
b 16 AAPTSTCA 23

RESULT 13  
US-09-765-272-170  
Sequence 170, Application US/09765272  
Patent No. US20020061545A1  
GENERAL INFORMATION:  
APPLICANT: Choi et al.  
TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines  
NUMBER OF SEQUENCES: 452  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/765.272  
FILING DATE: 22-Jan-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/961.083  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Brookes, A. Anders  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PB340P2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 170:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 247 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 170:  
US-09-765-272-170

Query Match 2.2% Score 7; DB 10; Length 247;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 252 ELVDAKG 258  
b 139 ELVDAKG 145

RESULT 14  
US-09-738-626-3805  
Sequence 3805, Application US/09738626  
Publication No. US20020197605A1  
GENERAL INFORMATION:  
APPLICANT: NAKAGAWA, SATOSHI  
APPLICANT: MIZOGUCHI, HIROSHI  
APPLICANT: ANDO, SEIKO  
APPLICANT: HAYASHI, MIKIRO  
APPLICANT: OCHIAT, KEIKO  
APPLICANT: YOKOI, HARUHIKO  
APPLICANT: TATEISHI, NAOKO  
APPLICANT: SENOH, AKIHIRO  
APPLICANT: IKEDA, MASATO  
APPLICANT: OKAKI, AKIO  
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
FILE REFERENCE: 249-125  
CURRENT APPLICATION NUMBER: US/09/738.626  
CURRENT FILING DATE: 2000-12-18  
PRIOR APPLICATION NUMBER: JP 95/377484  
PRIOR FILING DATE: 1999-12-16  
PRIOR APPLICATION NUMBER: JP 00/159162  
PRIOR FILING DATE: 2000-04-07  
PRIOR APPLICATION NUMBER: JP 00/280908  
PRIOR FILING DATE: 2000-08-03  
NUMBER OF SEQ ID NOS: 7059  
SOFTWARE: PatentIn ver. 3.0  
SEQ ID NO 3805  
LENGTH: 489  
TYPE: PRT  
ORGANISM: Corynebacterium glutamicum  
US-09-738-626-3805

Query Match 2.2% Score 7; DB 9; Length 489;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 11 LIANLIA 17  
b 198 LIANLIA 204

RESULT 15  
US-09-878-672-2  
Sequence 2, Application US/09878672  
Publication No. US20030049812A1  
GENERAL INFORMATION:  
APPLICANT: Wallis, Nicola G.  
TITLE OF INVENTION: NOVEL histidine kinase  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dechert Price & Rhoads  
STREET: 4000 Bell Atlantic Tower, 1717 Arch St  
CITY: Philadelphia  
STATE: PA  
COUNTRY: US  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/878.672  
FILING DATE: 11-Jun-2001  
CLASSIFICATION: <Unknown>

09978274  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/916,481  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Dickinson, Todd Q  
REGISTRATION NUMBER: 28,354  
REFERENCE/DOCKET NUMBER: GM10022-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-994-2252  
TELEFAX: 215-994-2222  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 554 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
JS-09-878-672-2

Query Match 2.2%; Score 7; DB 9; Length 554;  
Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
2Y 252 ELVDKAG 258  
1111111  
2b 158 ELVDKAG 164

Search completed: May 28, 2003, 10:21:55  
Job time : 58 secs

SRNT

Printed: 11/17/2004





09978274  
CURRENT FILING DATE: 2001-10-15  
PRIOR APPLICATION NUMBER: 0025225.4  
PRIOR FILING DATE: 2000-10-14  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 19  
LENGTH: 1092  
TYPE: DNA  
ORGANISM: Unknown  
FEATURE:  
OTHER INFORMATION: PAP-S/Cystatin fusion  
US-09-978-274A-19

Query Match 75.3% Score 712; DB 10; Length 1092;  
Best Local Similarity 99.9% Pred. No. 0;  
Matches 762; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 73 ATAAATACGATCACCTTTGATGCTGGAATGCCACCATTAACAAATATGCCACCTTTATG 132  
DB 4 ATAAATACGATCACCTTTGATGCTGGAATGCCACCATTAACAAATATGCCACCTTTATG 63

QY 133 GAATCTCTTCGTAATCAACGGAAGATCCAAACTAAATGCTATGCCATACCAATGCTA 192  
DB 64 GAATCTCTTCGTAATCAACGGAAGATCCAAACTAAATGCTATGCCATACCAATGCTA 123

QY 193 CCTGATACCTAATTCGACCCCTAAGTACTATTGTTAAGCTCCAAAGTGCACCACTTAAA 252  
DB 124 CCTGATACCTAATTCGACCCCTAAGTACTATTGTTAAGCTCCAAAGTGCACCACTTAAA 183

QY 253 ACCATTACCTAATTCGACCCCTAAGTACTATTGTTAAGCTCCAAAGTGCACCACTTAAA 312  
DB 184 ACCATTACCTAATTCGACCCCTAAGTACTATTGTTAAGCTCCAAAGTGCACCACTTAAA 243

QY 313 AATGGCAATAGTCTGAGCAAGTAACTTATGATTAAGTAACTTAACTTAACTTAACTTAACT 372  
DB 244 AATGGCAATAGTCTGAGCAAGTAACTTATGATTAAGTAACTTAACTTAACTTAACTTAACT 303

QY 373 GTGGAGAACTACTCTTGTCTCAAGTAACTTATGATTAAGTAACTTAACTTAACTTAACTTAACT 432  
DB 304 GTGGAGAACTACTCTTGTCTCAAGTAACTTATGATTAAGTAACTTAACTTAACTTAACTTAACT 363

QY 433 ACCTTATATCCGACCATTCGAAAGTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 492  
DB 364 ACCTTATATCCGACCATTCGAAAGTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 423

QY 493 GGAATTCAAATAGTCTGAGCAAGTAACTTATGATTAAGTAACTTAACTTAACTTAACTTAACT 552  
DB 424 GGAATTCAAATAGTCTGAGCAAGTAACTTATGATTAAGTAACTTAACTTAACTTAACTTAACT 483

QY 553 AAACTGAGGCTTTTCTTCTAGTCTGAGCAAGTAACTTATGATTAAGTAACTTAACTTAACTTAACT 612  
DB 484 AAACTGAGGCTTTTCTTCTAGTCTGAGCAAGTAACTTATGATTAAGTAACTTAACTTAACTTAACT 543

QY 613 AAGTACATAGCAACCACTCAAGTAACTTATGATTAAGTAACTTAACTTAACTTAACTTAACTTAACT 672  
DB 544 AAGTACATAGCAACCACTCAAGTAACTTATGATTAAGTAACTTAACTTAACTTAACTTAACTTAACT 603

QY 673 GTAATTAATTTGGAGGAGTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 732  
DB 604 GTAATTAATTTGGAGGAGTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 663

QY 733 GGGGCTTTACCAACCACTTCAAGTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 792  
DB 664 GGGGCTTTACCAACCACTTCAAGTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 723

QY 793 AGAGTGGATGAATCAATCCTGATGCGCACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 835  
DB 724 AGAGTGGATGAATCAATCCTGATGCGCACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 766

Patent No. US20020116737A1  
GENERAL INFORMATION:  
APPLICANT: Thomas, Christopher  
APPLICANT: McPherson, Michael  
APPLICANT: Atkinson, Howard  
APPLICANT: Neelam, Anil  
TITLE OF INVENTION: PLANT CELL DEATH SYSTEM  
FILE REFERENCE: 9341-028  
CURRENT APPLICATION NUMBER: US/09/978,274A  
CURRENT FILING DATE: 2001-10-15  
PRIOR APPLICATION NUMBER: 0025225.4  
PRIOR FILING DATE: 2000-10-14  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 5  
LENGTH: 465  
TYPE: DNA  
ORGANISM: Physolacca americana  
US-09-978-274A-5

Query Match 48.5% Score 459; DB 10; Length 465;  
Best Local Similarity 100.0% Pred. No. 1.3e-237;  
Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 73 ATAAATACGATCACCTTTGATGCTGGAATGCCACCATTAACAAATATGCCACCTTTATG 132  
DB 4 ATAAATACGATCACCTTTGATGCTGGAATGCCACCATTAACAAATATGCCACCTTTATG 63

QY 133 GAATCTCTTCGTAATCAACGGAAGATCCAAACTAAATGCTATGCCATACCAATGCTA 192  
DB 64 GAATCTCTTCGTAATCAACGGAAGATCCAAACTAAATGCTATGCCATACCAATGCTA 123

QY 193 CCTGATACCTAATTCGACCCCTAAGTACTATTGTTAAGCTCCAAAGTGCACCACTTAAA 252  
DB 124 CCTGATACCTAATTCGACCCCTAAGTACTATTGTTAAGCTCCAAAGTGCACCACTTAAA 183

QY 253 ACCATTACCTAATTCGACCCCTAAGTACTATTGTTAAGCTCCAAAGTGCACCACTTAAA 312  
DB 184 ACCATTACCTAATTCGACCCCTAAGTACTATTGTTAAGCTCCAAAGTGCACCACTTAAA 243

QY 313 AATGGCAATAGTCTGAGCAAGTAACTTATGATTAAGTAACTTAACTTAACTTAACTTAACT 372  
DB 244 AATGGCAATAGTCTGAGCAAGTAACTTATGATTAAGTAACTTAACTTAACTTAACTTAACT 303

QY 373 GTGGAGAACTACTCTTGTCTCAAGTAACTTATGATTAAGTAACTTAACTTAACTTAACTTAACT 432  
DB 304 GTGGAGAACTACTCTTGTCTCAAGTAACTTATGATTAAGTAACTTAACTTAACTTAACTTAACT 363

QY 433 ACCTTATATCCGACCATTCGAAAGTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 492  
DB 364 ACCTTATATCCGACCATTCGAAAGTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 423

QY 493 GGAATTCAAATAGTCTGAGCAAGTAACTTATGATTAAGTAACTTAACTTAACTTAACTTAACT 531  
DB 424 GGAATTCAAATAGTCTGAGCAAGTAACTTATGATTAAGTAACTTAACTTAACTTAACTTAACT 462

RESULT 5  
US-09-978-274A-7  
Sequence 7, Application US/09978274A  
Patent No. US20020116737A1  
GENERAL INFORMATION:  
APPLICANT: Thomas, Christopher  
APPLICANT: McPherson, Michael  
APPLICANT: Atkinson, Howard  
APPLICANT: Neelam, Anil  
TITLE OF INVENTION: PLANT CELL DEATH SYSTEM  
FILE REFERENCE: 9341-028  
CURRENT APPLICATION NUMBER: US/09/978,274A  
CURRENT FILING DATE: 2001-10-15  
PRIOR APPLICATION NUMBER: 0025225.4  
PRIOR FILING DATE: 2000-10-14  
NUMBER OF SEQ ID NOS: 32



SOFTWARE: PatentIn version 3.1  
SEQ ID NO 7  
LENGTH: 333  
TYPE: DNA  
ORGANISM: Phytolacca americana  
us-09-978-274A-7

Query Match 34.8% Score 329; DB 10; Length 333;  
Best Local Similarity 100.0%; Pred. No. 2.3e-167;  
Matches 329; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 532 GGAGTTGATTCATCCCTGTAAAACTGAGCGCTTTTCTACTGGTAGCCATCCAAATG 591  
Db 4 GGAGTTGATTCATCCCTGTAAAACTGAGCGCTTTTCTACTGGTAGCCATCCAAATG 63

QY 592 GTTTCAGAGCGAGCGGATTCAGTACATAGAGCAACCACTCAAGACTAATTTTATAGA 651  
Db 64 GTTTCAGAGCGAGCGGATTCAGTACATAGAGCAACCACTCAAGACTAATTTTATAGA 123

QY 652 GCATTTCACCTGATCCCAAGTAATTAATTGAGGAGCAAGTGGGCAAAATCTCTGAG 711  
Db 124 GCATTTCACCTGATCCCAAGTAATTAATTGAGGAGCAAGTGGGCAAAATCTCTGAG 183

QY 712 GCAATTCACAATCCCAAGTAATGGGCTTTTACCCAAACCCTTACCTAGCTAGTGGATGCAAA 771  
Db 184 GCAATTCACAATCCCAAGTAATGGGCTTTTACCCAAACCCTTACCTAGCTAGTGGATGCAAA 243

QY 772 GGTACCAAGTGGATAGTCTTACAGTGGATGAATCAATCGTATGTGGCACTCCTTAAG 831  
Db 244 GGTACCAAGTGGATAGTCTTACAGTGGATGAATCAATCGTATGTGGCACTCCTTAAG 303

QY 832 TAGCTTAATGGAACCTGTCAGACAACCTTA 860  
Db 304 TAGCTTAATGGAACCTGTCAGACAACCTTA 332

RESULT 6  
US-09-978-274A-15  
Sequence 15, Application US/09978274A  
Patent No. US20020116737A1  
GENERAL INFORMATION:  
APPLICANT: Thomas, Christopher  
APPLICANT: McPherson, Michael  
APPLICANT: Atkinson, Howard  
APPLICANT: Neelam, Anil  
TITLE OF INVENTION: PLANT CELL DEATH SYSTEM  
FILE REFERENCE: 9341-028  
CURRENT FILING DATE: 2001-10-15  
PRIOR APPLICATION NUMBER: 0025225.4  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 15  
LENGTH: 42  
TYPE: DNA  
ORGANISM: Unknown  
FEATURE:  
OTHER INFORMATION: PSXDF primer  
US-09-978-274A-15

Query Match 4.4% Score 42; DB 10; Length 42;  
Best Local Similarity 100.0%; Pred. No. 2.7e-12;  
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 736 GCTTTACCAACCACTGAGCTAGTGGATGCAAGGTACC 777  
Db 1 GCTTTACCAACCACTGAGCTAGTGGATGCAAGGTACC 42

RESULT 7  
us-09-978-274A-16/c  
Sequence 16, Application US/09978274A

Patent No. US20020116737A1  
GENERAL INFORMATION:  
APPLICANT: Thomas, Christopher  
APPLICANT: McPherson, Michael  
APPLICANT: Atkinson, Howard  
APPLICANT: Neelam, Anil  
TITLE OF INVENTION: PLANT CELL DEATH SYSTEM  
FILE REFERENCE: 9341-028  
CURRENT APPLICATION NUMBER: US/09/978,274A  
CURRENT FILING DATE: 2001-10-15  
PRIOR APPLICATION NUMBER: 0025225.4  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 16  
LENGTH: 42  
TYPE: DNA  
ORGANISM: Unknown  
FEATURE:  
OTHER INFORMATION: PSXDR primer  
US-09-978-274A-16

Query Match 4.4% Score 42; DB 10; Length 42;  
Best Local Similarity 100.0%; Pred. No. 2.7e-12;  
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 735 GCCTTTACCAACCACTGAGCTAGTGGATGCAAGGTACC 776  
Db 42 GCCTTTACCAACCACTGAGCTAGTGGATGCAAGGTACC 1

RESULT 8  
US-09-978-274A-13  
Sequence 13, Application US/09978274A  
Patent No. US20020116737A1  
GENERAL INFORMATION:  
APPLICANT: Thomas, Christopher  
APPLICANT: McPherson, Michael  
APPLICANT: Atkinson, Howard  
APPLICANT: Neelam, Anil  
TITLE OF INVENTION: PLANT CELL DEATH SYSTEM  
FILE REFERENCE: 9341-028  
CURRENT APPLICATION NUMBER: US/09/978,274A  
CURRENT FILING DATE: 2001-10-15  
PRIOR APPLICATION NUMBER: 0025225.4  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 13  
LENGTH: 52  
TYPE: DNA  
ORGANISM: Unknown  
FEATURE:  
OTHER INFORMATION: PS2BF primer  
US-09-978-274A-13

Query Match 3.2% Score 30; DB 10; Length 52;  
Best Local Similarity 100.0%; Pred. No. 8.5e-06;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 532 GGAGTTGATTCATTCCTGTAAAACTGAG 561  
Db 23 GGAGTTGATTCATTCCTGTAAAACTGAG 52

RESULT 9  
US-09-978-274A-31  
Sequence 31, Application US/09978274A  
Patent No. US20020116737A1  
GENERAL INFORMATION:  
APPLICANT: Thomas, Christopher  
APPLICANT: McPherson, Michael  
APPLICANT: Atkinson, Howard

APPLICANT: Neelam, Anil  
 TITLE OF INVENTION: PLANT CELL DEATH SYSTEM  
 FILE REFERENCE: 9341-028  
 CURRENT APPLICATION NUMBER: US/09/978,274A  
 CURRENT FILING DATE: 2001-10-15  
 PRIOR APPLICATION NUMBER: 0025225.4  
 PRIOR FILING DATE: 2000-10-14  
 NUMBER OF SEQ ID NOS: 32  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 31  
 LENGTH: 1368  
 TYPE: DNA  
 ORGANISM: Phytolacca americana  
 US-09-978-274A-31

Query Match 3.1%; Score 29; DB 10; Length 1368;  
 Best Local Similarity 100.0%; Pred. No. 3.9e-05;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 915 TAATCTTGTCATCTATTGAAGGATTCT 943  
 DB 1136 TAATCTTGTCATCTATTGAAGGATTCT 1164

RESULT 10  
 US-09-978-274A-30  
 Sequence 30, Application US/09978274A  
 Patent No. US20020116737A1  
 GENERAL INFORMATION:  
 APPLICANT: Thomas, Christopher  
 APPLICANT: McPherson, Michael  
 APPLICANT: Atkinson, Howard  
 APPLICANT: Neelam, Anil  
 TITLE OF INVENTION: PLANT CELL DEATH SYSTEM  
 CURRENT APPLICATION NUMBER: US/09/978,274A  
 CURRENT FILING DATE: 2001-10-15  
 PRIOR APPLICATION NUMBER: 0025225.4  
 PRIOR FILING DATE: 2000-10-14  
 NUMBER OF SEQ ID NOS: 32  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 30  
 LENGTH: 1379  
 TYPE: DNA  
 ORGANISM: Phytolacca americana  
 US-09-978-274A-30

Query Match 3.1%; Score 29; DB 10; Length 1379;  
 Best Local Similarity 100.0%; Pred. No. 3.9e-05;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 915 TAATCTTGTCATCTATTGAAGGATTCT 943  
 DB 1136 TAATCTTGTCATCTATTGAAGGATTCT 1164

RESULT 11  
 US-09-978-274A-12/c  
 Sequence 12, Application US/09978274A  
 Patent No. US20020116737A1  
 GENERAL INFORMATION:  
 APPLICANT: Thomas, Christopher  
 APPLICANT: McPherson, Michael  
 APPLICANT: Atkinson, Howard  
 APPLICANT: Neelam, Anil  
 TITLE OF INVENTION: PLANT CELL DEATH SYSTEM  
 CURRENT APPLICATION NUMBER: US/09/978,274A  
 CURRENT FILING DATE: 2001-10-15  
 PRIOR APPLICATION NUMBER: 0025225.4  
 PRIOR FILING DATE: 2000-10-14  
 NUMBER OF SEQ ID NOS: 32  
 SOFTWARE: PatentIn version 3.1

SEQ ID NO 12  
 LENGTH: 47  
 TYPE: DNA  
 ORGANISM: Unknown  
 FEATURE:  
 OTHER INFORMATION: PS2SR primer  
 US-09-978-274A-12

Query Match 2.9%; Score 27; DB 10; Length 47;  
 Best Local Similarity 100.0%; Pred. No. 0.00035;  
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 834 GGTAAATGGAACCTGTTCACACACACTTA 860  
 DB 47 CGTAAATGGAACCTGTTCACACACACTTA 21

RESULT 12  
 US-09-978-274A-14/c  
 Sequence 14, Application US/09978274A  
 Patent No. US20020116737A1  
 GENERAL INFORMATION:  
 APPLICANT: Thomas, Christopher  
 APPLICANT: McPherson, Michael  
 APPLICANT: Atkinson, Howard  
 APPLICANT: Neelam, Anil  
 TITLE OF INVENTION: PLANT CELL DEATH SYSTEM  
 CURRENT APPLICATION NUMBER: US/09/978,274A  
 CURRENT FILING DATE: 2001-10-15  
 PRIOR APPLICATION NUMBER: 0025225.4  
 PRIOR FILING DATE: 2000-10-14  
 NUMBER OF SEQ ID NOS: 32  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 14  
 LENGTH: 49  
 TYPE: DNA  
 ORGANISM: Unknown  
 FEATURE:  
 OTHER INFORMATION: PS1SR primer  
 US-09-978-274A-14

Query Match 2.9%; Score 27; DB 10; Length 49;  
 Best Local Similarity 100.0%; Pred. No. 0.00035;  
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 505 CTCACAGTCGACATTCGAAATCTCT 531  
 DB 49 CTCACAGTCGACATTCGAAATCTCT 23

RESULT 13  
 US-09-978-274A-11  
 Sequence 11, Application US/09978274A  
 Patent No. US20020116737A1  
 GENERAL INFORMATION:  
 APPLICANT: Thomas, Christopher  
 APPLICANT: McPherson, Michael  
 APPLICANT: Atkinson, Howard  
 APPLICANT: Neelam, Anil  
 TITLE OF INVENTION: PLANT CELL DEATH SYSTEM  
 CURRENT APPLICATION NUMBER: US/09/978,274A  
 CURRENT FILING DATE: 2001-10-15  
 PRIOR APPLICATION NUMBER: 0025225.4  
 PRIOR FILING DATE: 2000-10-14  
 NUMBER OF SEQ ID NOS: 32  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 11  
 LENGTH: 48  
 TYPE: DNA  
 ORGANISM: Unknown  
 FEATURE:

OTHER INFORMATION: PSIBF primer

US-09-978-274A-11

Query Match 2.8%; Score 26; DB 10; Length 48;  
Best Local Similarity 100.0%; Pred. No. 0.0012;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 73 ATAAATACGATCACCTTTGATGCTGG 98  
Db 23 ATAAATACGATCACCTTTGATGCTGG 48

RESULT 14

US-09-978-274A-20/c

; Sequence 20, Application US/09978274A  
; Patent No. US20020116737A1  
; GENERAL INFORMATION:  
; APPLICANT: Thomas, Christopher  
; APPLICANT: McPherson, Michael  
; APPLICANT: Atkinson, Howard  
; APPLICANT: Neelam, Anil  
; TITLE OF INVENTION: PLANT CELL DEATH SYSTEM  
; FILE REFERENCE: 9341-028  
; CURRENT APPLICATION NUMBER: US/09/978,274A  
; PRIOR FILING DATE: 2001-10-15  
; PRIOR APPLICATION NUMBER: 0025225.4  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 20  
; LENGTH: 45  
; TYPE: DNA

ORGANISM: Unknown

FEATURE:

OTHER INFORMATION: PCS-PAPSR primer  
US-09-978-274A-20

Query Match 2.6%; Score 25; DB 10; Length 45;  
Best Local Similarity 100.0%; Pred. No. 0.0042;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 811 CGTGATGTGGCACTCCTTAAGTACG 835  
Db 45 CGTGATGTGGCACTCCTTAAGTACG 21

RESULT 15

US-09-978-274A-9

; Sequence 9, Application US/09978274A  
; Patent No. US20020116737A1  
; GENERAL INFORMATION:  
; APPLICANT: Thomas, Christopher  
; APPLICANT: McPherson, Michael  
; APPLICANT: Atkinson, Howard  
; APPLICANT: Neelam, Anil  
; TITLE OF INVENTION: PLANT CELL DEATH SYSTEM  
; FILE REFERENCE: 9341-028  
; CURRENT APPLICATION NUMBER: US/09/978,274A  
; PRIOR FILING DATE: 2001-10-15  
; PRIOR APPLICATION NUMBER: 0025225.4  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 9  
; LENGTH: 43  
; TYPE: DNA

ORGANISM: Unknown

FEATURE:

OTHER INFORMATION: PPSIBF primer  
US-09-978-274A-9

Query Match 2.5%; Score 24; DB 10; Length 43;  
Best Local Similarity 100.0%; Pred. No. 0.015;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGAAGGTGATGCTTGTAGTTGTG 24  
Db 20 ATGAAGGTGATGCTTGTAGTTGTG 43

Search completed: June 2, 2003, 05:53:35  
Job time : 350 secs